

Experimentelle Studien zur schlafabhängigen Konsolidierung großmotorischer Bewegungsfolgen

Kumulative Dissertation
zur Erlangung des Grades eines
Doktors der Philosophie
der Fakultät für Empirische Humanwissenschaften und Wirtschaftswissenschaft
der Universität des Saarlandes

vorgelegt von
Andreas Malangré
Matrikelnummer: 250105
Aus: Mainz
Saarbrücken, 2020

Dekan:
Univ.-Prof. Dr. Stefan Strohmeier

Berichterstatter:
Prof. Dr. Klaus Blischke
Univ.-Prof. Dr. Stefan Panzer

Tag der Disputation: 26 .11.2019

Inhaltsverzeichnis

EXPERIMENTELLE STUDIEN ZUR SCHLAFABHÄNGIGEN KONSOLIDIERUNG GROßMOTORISCHER BEWEGUNGSFOLGEN	0
INHALTSVERZEICHNIS	2
SCHLAF UND DIE KONSOLIDIERUNG NEU GELERNTER MOTORISCHER AUFGABEN.....	5
1 EINLEITUNG	5
2 THEORETISCHE GRUNDLAGEN: GEDÄCHTNISKONSOLIDIERUNG & SCHLAF	6
2.1 Konsolidierung.....	6
2.2 Schlaf.....	9
Einteilung in Schlafphasen (Sleepstages).....	10
NREM-Schlaf.....	10
REM-Schlaf.....	11
2.3 Verbindung von Schlaf- und Wachheit mit Konsolidierungsprozessen	11
2.4 Schlaf und die Konsolidierung motorischer Aufgaben.....	12
2.5 Von Finger- zu Armbewegungsaufgaben: Unterschiedliche Befundlagen bei großmotorischen Aufgaben.....	14
2.6 Weitere Untersuchungen alltagsnaher Aufgaben.....	16
2.7 Begriffsklärung: Verwendung von Terminologien in dieser Arbeit.....	19
2.7.1 Großmotorisch	19
2.7.2 Komplexität	20
3 VORGEHENSWEISE FÜR DIE PUBLIKATIONEN DER VORLIEGENDEN ARBEIT	23
3.1 Paradigma, Messapparat, Versuchsaufbau und experimentelles Protokoll für die vorliegenden Publikationen.....	26
3.1.1 Das Paradigma.....	26
3.1.2 Ablauf der Experimente	27
3.1.3 Versuchsaufbau	27
3.1.4 Experimentelles Protokoll	30
3.1.5 Abhängige Variablen	31
3.1.6 Zwei Komponenten für die Identifikation von leistungssteigernden Konsolidierungseffekten.....	32
4 ÜBERBLICK ÜBER DIE PUBLIKATIONEN	34

<i>4.1 Publikation 1: Sleep-related offline learning in a complex arm movement sequence (Malangré, Leinen & Blischke, 2014)</i>	34
4.1.1 Ziel	34
4.1.2 Aufbau	34
4.1.3 Ergebnisse	35
<i>4.2 Publikation 2: Chunk Concatenation evolves with practice and sleep-related enhancement consolidation in a complex arm movement sequence (Blischke & Malangré, 2016)</i>	35
4.2.1 Ziel	35
4.2.2 Theoretische und empirische Ausgangslage.....	36
4.2.3 Methode	38
4.2.4 Ergebnisse.....	39
<i>4.3 Publikation 3: Sleep-related offline improvements in gross motor task performance occur under free recall requirements (Malangré & Blischke, 2016)</i>	40
4.3.1 Ziel	40
4.3.2 Methode	42
4.3.3 Ergebnisse.....	44
<i>4.4 Publikation 4: Task complexity modulates sleep-related offline learning in sequential motor skills (Blischke & Malangré, 2017)</i>	46
4.4.1 Ziel	46
4.4.2 Methode	47
4.4.3 Ergebnisse.....	49
4.4.4 Weiterführende Überlegungen und Auswertungsschritte	49
4.5 ZUSAMMENFASSUNG DER VIER PUBLIKATIONEN	51
5 EINIGE ÜBERLEGUNGEN ZUR AUSWERTUNGSMETHODE	55
5.1 Methodische Fragestellungen	56
5.1.1 Datenmittlung	56
5.1.2 Warm-up-Decrements	58
5.1.3 Reaktive Hemmung.....	64
5.1.4 Early Performance Boost und reaktive Hemmung?.....	68
5.1.5 Ermüdung.....	69
5.1.6 Mit Übungsende erreichtes Leistungsniveau und schlafbasierte Leistungssteigerungen.....	72
5.2 Fazit: Keine Verzerrung der statistischen Auswertung bei der Peg-Plugging-Aufgabe durch Datenmittlung von Leistungsdaten in den hier vorgelegten Untersuchungen.....	73

5.3 Weitere Aspekte für künftige Untersuchungen	76
5.3.1 Time of Testing	76
5.3.2 Naps	77
5.3.3 Type of Task	77
5.3.4 Testing-Effekt	78
5.3.5 Alter und schlafbasierte Leistungsveränderung	79
5.3.6 Weitere Komplexitätsfaktoren.....	82
6 FAZIT.....	84
7 LITERATURVERZEICHNIS.....	85
DANKSAGUNG.....	93
VERZEICHNIS DISSERTATIONSRELEVANTER PUBLIKATIONEN	94
WEITERE PUBLIKATIONEN	94
VORTRÄGE UND KONFERENZBEITRÄGE MIT REVIEW DES ABSTRACTS	95

Schlaf und die Konsolidierung neu gelernter motorischer Aufgaben.

1 Einleitung

Neue Inhalte gelangen auf verschiedensten Wegen in das menschliche Gehirn und werden dort zu dauerhaftem Gedächtnisbesitz. Für das Erlernen motorischer Aufgaben ist z.B. unter anderem das wiederholte aktive Ausführen der entsprechenden Aufgabe erforderlich.

Neben dieser aktiven Komponente ist jedoch eine weitere Verarbeitung der neu aufgenommenen Gedächtnisinhalte vonnöten, damit diese von einem initial labilen Status in einen stabileren, dauerhaft abrufbaren Zustand überführt werden können (McGaugh, 2000). Für diese Überführung sorgen sogenannte Konsolidierungsprozesse. Sie tragen Sorge dafür, dass eine einmal gelernte Tätigkeit auf einem mehr oder weniger stabilen Leistungsniveau in der Zukunft wieder aufgenommen werden kann. In den letzten 20 Jahren hat sich die Erkenntnis durchgesetzt, dass Konsolidierungsprozesse neben dem Verstreichen von Zeit auch in Abhängigkeit von Prozessen, die unterschiedlichen Aktivitätszuständen des Gehirns (Wachheit vs. Schlaf) zugeordnet werden können, auftreten.

Der Annahme, dass Schlaf – im Gegensatz zum alleinigen Verstreichen von Zeit im Wachzustand – eine entscheidende Rolle für die Konsolidierung neu gelernter *motorischer* Aufgaben zukommt, wurde seit den 1990er Jahren verstärkte Aufmerksamkeit geschenkt (Plihal & Born, 1997; Walker et al., 2002; Walker, 2003, Fischer et al., 2002, 2005).

Wurden anfangs vor allem Fingerbewegungssequenzen für die Untersuchung von Konsolidierungseffekten genutzt, hat die Forschung sich in den letzten zehn Jahren auch der Bedeutung des Schlafes für die Konsolidierung *großmotorischer* Aufgaben gewidmet (Blischke et al., 2008; Kempner & Richmond, 2012; Genzel et al., 2015).

2 Theoretische Grundlagen: Gedächtniskonsolidierung & Schlaf

Die Anfänge der Forschung zur Konsolidierung von Gedächtnisinhalten datieren ebenso wie die ersten Schritte der Schlafforschung zurück bis zum Ende des 19. bzw. zum Beginn des 20. Jahrhunderts (McGaugh, 2000; Patrick & Gilbert, 1896). Die Bedeutung des Schlafes für die Konsolidierung von Gedächtnisinhalten wurde ebenfalls schon zu Beginn des 20. Jahrhunderts untersucht (siehe bspw. Dahl, 1928). Bei Untersuchungen des „Wiedererkennens“ sinnloser Silben und Flächenfiguren stellte Dahl (1928) fest: „In allen Fällen war nach dem Schlafen weniger vergessen worden als nach dem entsprechenden Wachintervall.“ (Dahl, 1928).

2.1 Konsolidierung

Der Begriff „Konsolidierung“ wird in der Literatur in Verbindung mit einer Reihe von Phänomenen verwendet, bspw. der Überführung von „memories“ – Gedächtnisinhalten – aus einem initial fragilen Status in eine dauerhaft abrufbare Information (McGaugh, 2000; Robertson & Cohen, 2006). Damit wird sichergestellt, dass relevante Erinnerungen längerfristig verfügbar bleiben und nicht von dem steten Fluss an neuen Informationen überdeckt werden (Born & Wilhelm, 2012). Die Konsolidierung nach dem Neulernen erfordert dabei vor allem das Verstreichen von Zeit, ohne dass weitere, konkurrierende Lernprozesse stattfinden (Müller & Blischke, 2009). Dabei spielen verschiedene Hirnareale (bspw. Hippocampus, Amygdala, Basalganglien und verschiedene Areale im Neocortex) sowie Prozesse auf hormoneller, synaptischer und molekularer Ebene eine Rolle (McGaugh, 2000). Während Erinnerungen im Kurzzeitgedächtnis extrem schnell gespeichert (und daraufhin schnell wieder „vergessen“) werden, benötigt der Aufbau einer langfristig verfügbaren Gedächtnisrepräsentation deutlich mehr Zeit (McGaugh, 2000).

In weiteren Literaturstellen findet sich der Konsolidierungsbegriff in Verbindung mit extern beobachtbaren Verhaltenskriterien wie der Behaltensleistung oder der Resistenz gegen Interferenz nach einem Behaltensintervall (Doyon & Benali, 2005). Walker (2005) spricht von einem zweiphasigen Konsolidierungsmodell, welches den

Konsolidierungsprozess in eine Phase der Stabilisierung und eine Phase der Leistungssteigerung unterteilt. Dabei scheint sich, zumindest für den Autor dieser Synopse, stillschweigend auf den oben erwähnten Konsolidierungsbegriff nach McGaugh (2000) bezogen zu werden, denn eine genauere Definition des Begriffes an sich findet nicht statt.

Studien, welche die Resistenz einer neu erworbenen Repräsentation gegen Störungen untersuchen – bspw. durch die unmittelbar folgende Aneignung weiterer, ähnlicher Gedächtnisinhalte –, geben Aufschluss über den zeitlichen Verlauf der Konsolidierung. So fanden Brashers-Krug et al. (1996), dass bei einer neu gelernten Zielbewegungsaufgabe die Leistung beim Abruf am Tag nach der Aneignungsphase deutlich schlechter ist, wenn eine zweite Zielbewegungsaufgabe direkt im Anschluss an die Übungsphase der Erstaufgabe gelernt wird. Genauso verhält es sich für das Erlernen der zweiten Aufgabe fünf Minuten und eine Stunde nach Beendigung der Aneignungsphase der ersten Aufgabe. Erst wenn der Abstand zwischen der Aneignung beider Aufgaben vier Stunden und mehr beträgt, gelingt die Ausführung am nächsten Tag störungsfrei (Brashers-Krug et al., 1996). Diese Befunde deuten darauf hin, dass die ersten vier Stunden nach dem Erlernen einer motorischen Aufgabe entscheidend für die Stabilisierung der Gedächtnisrepräsentation sind (siehe dazu auch Nettersheim et al., 2015). Weitere Untersuchungen (Shadmehr & Holcomb, 1999) haben gezeigt, dass bei kurzem Zeitabstand zwischen dem Erlernen zweier ähnlicher Adaptationsaufgaben während der Aneignung der später zu lernenden Aufgabe die gleichen Hirnregionen aktiv sind, die auch während der Aneignungsphase der ersten Aufgabe aktiv waren. Dies ist allerdings nicht mehr der Fall, wenn mehr als 5,5 Stunden zwischen der Aneignung beider Aufgaben lagen. Darüber hinaus zeigt sich nunmehr eine verstärkte Aktivität im ventralen präfrontalen Kortex, der eine hemmende Funktion zukommt, die ihre Wirkung über die Basalganglien entfaltet. Diese Beobachtung wird dahingehend interpretiert, dass die interne Repräsentation der zuerst gelernten Aufgabe nach einer Zeitdauer von 5,5 Stunden gehemmt wird und somit die Ausbildung einer zweiten, neuen Repräsentation erlaubt (Blischke & Müller, 2009).

Die Auswahl von Gedächtnisinhalten für die folgende Konsolidierung ist zudem abhängig von weiteren Faktoren, bspw. dem emotionalen Zustand zum Zeitpunkt der Encodierung (McGaugh et al., 2000) oder der subjektiven Relevanz einer Erinnerung für die Zukunft (Born & Wilhelm, 2012).

Erweiterungen der Konsolidierungsdefinition

Walker et al. (2004) schlagen eine Erweiterung der Konsolidierungsdefinition vor: So umfasst der Begriff „Konsolidierung“ nicht nur die Stabilisierung von Gedächtnisinhalten, sondern kann auch eine Optimierung dieser Inhalte umfassen. Zudem beinhaltet der Begriff „Konsolidierung“ mehrere Phasen der Verarbeitung von Gedächtnisinhalten im Anschluss an deren Encodierung. Dabei treten die Verarbeitungsphasen in Abhängigkeit von spezifischen Hirnzuständen, wie Wachheit oder Schlaf, auf.

Blischke & Müller (2009) unterscheiden zwischen den Begriffen „Konsolidierung“ und „Optimierung“. Konsolidierung bezeichnet dabei die Stabilisierung neu erworbener kognitiver Repräsentationen ohne weitere Aufgabenausführung und erfordert Ruhe. Mit „Ruhe“ ist dabei gemeint, dass erstens keinerlei weitere Ausführung der Aufgabe stattfindet und zweitens keine weitere, ähnliche Aufgabe erlernt oder ausgeführt wird. Das Erlernen neuer Aktivitäten im Ruhezeitraum führt zur Störung der Konsolidierung und damit zu Interferenzeffekten.

Die „Optimierung“ einer neu erworbenen Repräsentation erfordert laut Blischke & Müller (2009) dagegen i.d.R. Schlaf und führt zu einer Leistungssteigerung bei der Kriteriumsleistungsaufgabe.

Zur Verwendung der Terminologie in dieser Arbeit

In dieser Arbeit wird der Konsolidierungsbegriff in Anlehnung an die Definition von McGaugh (2000) verwendet: die Überführung von Gedächtnisinhalten aus einem initial fragilen Zustand in eine dauerhaft abrufbare Erinnerung. Damit ist die Verwendung des Konsolidierungsbegriffes unabhängig von verhaltensseitig beobachtbaren Veränderungen der Kriteriumsleistung und beinhaltet alle Prozesse, die dazu beitragen, dass ein Gedächtnisinhalt in einen stabileren, dauerhaft abrufbaren Zustand überführt wird.

Als *Konsolidierungsprozesse* werden die dieser Überführung zugrundeliegenden (hirn-)physiologischen Prozesse bezeichnet.

Sind Leistungsveränderungen nach einem Behaltensintervall zu beobachten, während dessen Probanden vorrangig geschlafen haben, wird von *schlafabhängigen Leistungsveränderungen* gesprochen.

Die dahinterstehende Annahme ist, dass diese *schlafbasierten* Leistungsveränderungen aufgrund von Konsolidierungsprozessen entstehen, die ausschließlich während des Schlafes stattfinden können.

Daneben laufen Konsolidierungsprozesse auch im Wachzustand ab und auch hier sind nach einem Behaltensintervall Leistungsveränderungen (Verbesserungen oder Verschlechterungen) beobachtbar. Bei solchen Veränderungen wird von zeitbasierten Leistungsveränderungen gesprochen. In der Regel fallen Leistungsveränderungen nach einem Wachintervall allerdings deutlich weniger stark aus (bspw. Walker et al., 2002).

2.2 Schlaf

Da sich diese Arbeit mit der *schlafbasierten* Konsolidierung von Gedächtnisinhalten befasst, soll in diesem Abschnitt ein kurzer Überblick über den Schlaf und seine Besonderheiten gegeben werden.

Born & Wilhelm (2011) zufolge ist Schlaf ein Zustand, der sich insbesondere durch die Abwesenheit von Bewusstsein vom Wachzustand unterscheidet. Welche Funktionen dem Schlaf genau zukommen, ist immer noch nicht sicher umfassend geklärt, allerdings existieren einige Hinweise darauf. Pilcher & Huffcutt (1996) zeigten, dass Schlafdeprivation zu Einschränkungen der Funktionsweise des menschlichen Körpers auf unterschiedliche Weise führt: So verschlechtern sich sowohl die Stimmung als auch motorische und kognitive Funktionen. Einige Forschungsgruppen (Born & Wilhelm, 2011; Genzel & Robertson, 2015) vermuten, dass die Überführung von Gedächtnisinhalten in das Langzeitgedächtnis nicht synchron mit der konstanten Verarbeitung neuer Eindrücke im Kurzzeitgedächtnis (während Wachheit) ablaufen kann. Daher liegt laut Born & Wilhelm (2011, p. 192) die Vermutung nahe, dass die Konsolidierung von Gedächtnisinhalten im Langzeitgedächtnis einen Zustand der Abwesenheit von Bewusstsein erfordert, den nur der Schlaf liefert.

Darüber hinaus könnte das Vorhandensein von Schlaf für allgemeine sowie spezielle Regenerationsprozesse vonnöten sein. Tononi & Cirelli (2014) vermuten bspw., dass nur Prozesse während des Schlafes vor einer synaptischen Überlastung schützen können und dafür sorgen, dass eine sogenannte synaptische Homöostase wiederhergestellt wird.

Tononi & Cirelli (2014) gehen davon aus, dass dem Schlaf aufgrund seiner alle Spezies übergreifenden Eigenschaft (Unterbrechung der Interaktionsmöglichkeit zwischen Umwelt, Körper und Gehirn) eine bedeutende Funktion zukommt. Da der Preis für die während Wachphasen stattfindenden synaptischen Prozesse recht hoch ist (bspw. erhöhter Energieverbrauch sowie eine Saturierung von Lernmöglichkeiten), könnte die Diskonnektivität mit der Umwelt während des Schlafes dazu dienen, eine Homöostase auf synaptischer Ebene wiederherzustellen (Tononi & Cirelli, 2014, p. 12). Dieser Prozess sorgt laut Tononi & Cirelli (2014, p. 14) dafür, dass während Wachheit wiederholt benutzte oder gut mit bisherigen Erinnerungen integrierte Synapsen erhalten werden. Im Gegensatz dazu könnten wenig genutzte synaptische Verbindungen während des Schlafes geschwächt werden („synaptic downscaling“).

Der Schlaf stellt dabei kein homogenes Phänomen dar, sondern erscheint nach (hirn-)physiologischen Zuständen und funktionalen Gesichtspunkten strukturiert.

Einteilung in Schlafphasen (Sleepstages)

Der menschliche Schlaf wird in der Regel in fünf verschiedene Schlafphasen eingeteilt, die durch unterschiedliche physiologische Phänomene charakterisiert werden.

In der gängigen Klassifizierung werden zwei sich grundsätzlich voneinander unterscheidende Arten von Schlaf unterschieden: Rapid-Eye-Movement(REM)- und Non-Rapid-Eye-Movement(NREM)-Schlaf lassen sich durch Aktivitätsmuster der Skelettmuskulatur (EMG-Aktivität), mit dem Elektroenzephalogramm erfassbare Hirnaktivität sowie Bewegung der Augen (Elektrookulogramm – EOG) voneinander unterscheiden (Walker, 2009). Während des Schlafes durchläuft der menschliche Organismus wiederholt ein sich ähnelndes Muster der verschiedenen Phasen von REM bis NREM-4. Dabei dominieren in der ersten Hälfte der Nacht die tieferen Schlafphasen NREM-3 und NREM-4. In der zweiten Hälfte der Nacht treten dagegen häufiger längere Phasen von REM- und NREM-2-Schlaf auf (Walker, 2005, p. 52).

NREM-Schlaf

NREM-Schlaf wird in vier weitere Phasen (NREM-1–4) unterteilt. Dabei steigt die Tiefe des Schlafes mit aufsteigender Nummerierung der Phasen (NREM-1: leichtester Schlaf, NREM-4: tiefster Schlaf). Die Phasen 3 und 4 werden zusammenfassend als Slow-Wave-

Sleep (SWS) bezeichnet. Je tiefer der Schlaf, desto niedriger die Frequenz der EEG-Aktivität im Gehirn (siehe Genzel et al., 2015; Walker, 2005, 2009).

NREM-1: EEG: Theta-Aktivität (4–8 Hz)

NREM-2: EEG: Theta-Aktivität, K-Komplexe (große, scharfe Ausschläge im EEG, Schlafspindeln (kurze synchronisierte Oszillationen mit 7–16 Hz) (siehe auch Walker, 2009; Genzel et al., 2015)

NREM-3: EEG: Delta-Aktivität: 0,5–4 Hz mit hoher Amplitude

NREM-4: EEG: Delta-Aktivität, synchrone Aktivität im Cortex

REM-Schlaf

REM-Phasen sind von erhöhter EEG-Aktivität bei gleichzeitiger Erschlaffung der Skelettmuskulatur gekennzeichnet und verlängern sich in der zweiten Schlafhälfte gegenüber der ersten Hälfte (Walker, 2005). Zusätzlich treten während des REM-Schlafes die namensgebenden „*Rapid Eye-Movements*“, schnelle horizontale Augenbewegungen in Schüben, auf.

Weiterhin lassen sich im Gegensatz zum NREM-Schlaf eher dem Wachzustand ähnliche (Genzel et al., 2014) desynchronisierte EEG-Aktivität im Gamma-Bereich (30–80 Hz) sowie das Auftreten sogenannter PGO-Wellen feststellen. Diese erhalten ihren Namen aufgrund des Auftretens in den Arealen *P*ons, Nucleus *g*eniculatus des Thalamus und im *occipitalen* Cortex (Walker & Robertson, 2004; Walker, 2005).

2.3 Verbindung von Schlaf- und Wachheit mit

Konsolidierungsprozessen

Die Idee, dass neben dem Verstreichen von Zeit sowie dem emotionalen Zustand zum Zeitpunkt der Encodierung (McGaugh, 2000, S. 248) auch der generelle Zustand (Wachheit vs. Schlaf) des Körpers während des Konsolidierungszeitraumes eine Rolle für die Konsolidierung von Gedächtnisinhalten spielen könnte, entstand bereits im frühen 20. Jahrhundert (Dahl, 1928; Ekstrand, 1967). Tilley & Empson (1978) brachten in einer Deprivationsstudie die Genauigkeit der Wiedergabe (Erinnerung) an eine vor dem Schlaf

gehörte Geschichte mit dem Durchlaufen verschiedener Schlafphasen in Verbindung. Speziell die Deprivation von REM-Schlaf führte zur Verringerung der Genauigkeit der erinnerten Geschichte. Im Gegensatz dazu hatte die Deprivation von NREM-4-Schlaf keine solche Effekte zur Folge (Tilley & Empson, 1978).

Bruce Ekstrand (1967) untersuchte, wie sich Schlaf auf die Erinnerung an Wortlisten auswirkt. Er fand heraus, dass sich die Probanden beim Vergleich der Behaltensleistung nach Schlaf und Wachheit besser nach dem durchgeschlafenen Intervall an vorher präsentierte Wortlisten erinnern konnten.

Die Untersuchung der Auswirkung des Schlafes auf die Verarbeitung kognitiver Gedächtnisinhalte hält bis heute an (Gais et al., 2006).

2.4 Schlaf und die Konsolidierung *motorischer* Aufgaben

In den vergangenen 15 Jahren haben sich Studien vermehrt auch der Untersuchung der Rolle des Schlafes für die Konsolidierung *motorischer* Aufgaben gewidmet (siehe auch King et al., 2017).

Dabei hat sich wiederholt gezeigt, dass die Konsolidierung neu gelernter Fingerbewegungssequenzen nach einem durchgeschlafenen Zeitraum nicht nur zur Stabilisierung der Kriteriumsleistung gegenüber der Leistung am Ende der Aneignungsphase führt. Auch wurde mehrfach gezeigt, dass sich die Kriteriumsleistung nach Schlaf verbessert. Im Gegensatz dazu führt das alleinige Verstreichen von Zeit im Wachzustand lediglich zu einer Stabilisierung der Kriteriumsleistung (Walker et al., 2003; Kuriyama et al., 2009). Der Schlaf scheint also nicht nur eine besondere Rolle für die Konsolidierung von kognitiven Gedächtnisinhalten zu spielen, sondern auch für die Konsolidierung von Gedächtnisinhalten nach dem Neulernen von motorischen Aufgaben.

Insbesondere die Arbeitsgruppen um Matthew Walker (Walker et al., 2002; Walker, 2003), Born (Born & Wilhelm, 2012; Wilhelm et al., 2013), Maquet et al. (2000, 2001) sowie Fischer et al. (2002) haben dazu beigetragen, ein besseres Verständnis der Konsolidierungsprozesse nach dem Neulernen motorischer Aufgaben zu erlangen.

Fast alle Studien dieser Gruppen (Walker, Stickgold, Kuriyama etc.) haben jedoch eines gemein: Sie nutzen stets eine ähnliche Form des gleichen experimentellen Paradigmas:

der Serial-Reaction-Time-Task (SRTT) bzw. der Finger-to-Thumb-Opposition-Task (Fischer et al., 2002).

Diese Aufgabe beinhaltet das serielle Wiederholen der jeweils gleichen Sequenz von fünf Tastendrücken auf einer Tastatur (bspw. Tasten 4-1-3-2-4 (Abbildung Walker et al., 2002) über einen Zeitraum von 30 Sekunden.

Dabei verharren in der Standardvariante vier Finger (kleiner Finger bis Zeigefinger) über der jeweiligen Taste, bis eine Aktion (Tastendruck) gefordert wird.

Diese Aufgabe lässt sich wie folgt charakterisieren:

- Serielle Reaktionszeitaufgabe
- Gelernt wird die Folge der Aktivierung verschiedener Endeffektoren.
- Kontinuierliche Aufgabenausführung ohne Pause nach Beendigung einer Sequenzausführung
- Aneignung unter expliziten Bedingungen: Der Lernende hat die bewusste Absicht, die Aufgabe zu erlernen, übt zielgerichtet und kann die Regelmäßigkeit der Aufgabe verbal ausdrücken (Müller & Blischke, 2009).
- Zumindest partiell deklarative Gedächtnisanforderungen
- Vier unilaterale Endeffektoren (Finger) involviert
- Geringe räumliche Präzisionsanforderungen¹
- Wenige biomechanische Freiheitsgrade involviert (d.h. pro Endeffektor ein Gelenk mit einer Bewegungsrichtung)
- Keine Aktivierung und Koordinierung großer Muskelgruppen erforderlich und keine Redundanzkontrolle erforderlich

Damit beinhaltet das SRTT-Paradigma sehr spezielle Anforderungen an die Bewegungskontrolle. Und obwohl Walker (2002) festhielt:

“These findings now add to a growing set of data suggesting that many, if not all, sensorimotor skills may require post-training sleep for the optimal consolidation.”

¹ Fitt's, räumliche Bewegung im Raum, Verortung der Ziele im allozentrischen oder egozentrischen Koordinatensystem.

(Walker et al. (2002), muss die Frage gestellt werden, ob aufgrund der genannten speziellen Aufgabencharakteristika eine solche Verallgemeinerung tatsächlich zulässig ist.

Bewegungsaufgaben in alltagsnahen Kontexten, wie dem Bewegungslernen im Sport oder der Rehabilitation, erfordern ein weit größeres Repertoire an Anforderungen, bspw. die Produktion diskreter Bewegungen (d.h. Richtungs- und Distanzkontrolle im dreidimensionalen Raum), hohe räumliche bzw. zeitliche Bewegungspräzision, die Kontrolle einzelner Gliedmaßen oder die Koordination mehrerer Gliedmaßen und damit andere Anforderungen an die Redundanzkontrolle. Diese Unterschiedlichkeit der motorischen Kontrollanforderungen lässt die Frage plausibel erscheinen, ob das Erlernen unterschiedlicher Aufgabenarten möglicherweise auch das Encodieren unterschiedlicher Aufgabenrepräsentationen zur Folge hat, die unterschiedliche Mechanismen der Konsolidierung erfordern.

2.5 Von Finger- zu Armbewegungsaufgaben: Unterschiedliche Befundlagen bei großmotorischen Aufgaben

Seit 2008 gehen Blischke et al. (2008) der Frage nach, ob auch bei anderen, zunehmend an alltäglichen Bewegungen orientierten Aufgaben ähnliche schlafbasierte Konsolidierungseffekte auftreten.

Zu diesem Zweck wurde eine „Strategie multipler Aufgaben“ (siehe Heuer, 1993) verfolgt. Sie sieht vor, aus einmal entdeckten Zusammenhängen Hypothesen aufzustellen und diese dann mit verschiedenen Aufgaben vom einfachen experimentellen Setting bis hin zu praxisnahen Aufgaben zu prüfen. Daneben gilt es, das Aufgabenspektrum auf derselben Komplexitätsebene von der Labor- bis zur Feldsituation zu prüfen.

Diese Strategie beinhaltet folgende Aufgaben und Schritte:

1. Replikation von Variationen der Serial Reaction Time Task mit Hinblick auf
 - a. Die Anzahl verwendeter Endeffektoren (Anzahl der Finger)
 - b. Die Anzahl der involvierten Extremitäten (Handtapping) (Schmidt et al., 2010)
 - c. Die Anzahl der Freiheitsgrade

2. Versuch der Replikation der Befunde mit weiteren Aufgaben, bspw.:

- a. Einer durch absolute und relative Zeitdauern und unterschiedliche räumliche Ziele definierte Handbewegungssequenz mit einem Endeffektor – unter Einbeziehung des ganzen Arms musste ein räumliches Ziel von einem Finger erreicht werden – mit vergleichsweise geringer Komplexität (vier klar voneinander unterscheidbare räumliche Ziele, drei zu kontrollierende Zeitintervalle, niedriger Fitts' ID) (kurz: Diamond-Tapping-Task – DTT)
- b. Einer ballistischen, großmotorischen Kraft-Parametrisierungs-Aufgabe unter Einbeziehung großer Muskelgruppen (Counter-Movement-Jump – CMJ)
- c. Einer visuomotorischen Verfolgungsaufgabe unter Einbeziehung eines Endeffektors
- d. Dem Peg-Plugging-Paradigma, einer komplexen, diskreten Sequenz-Produktionsaufgabe mit hohen räumlichen Präzisionsanforderungen

Das letztgenannte Peg-Plugging-Paradigma wurde auch im Rahmen der vorliegenden Arbeit genutzt.

Die Ergebnisse von Blischke et al. (2008) zeigen, dass die Replikation früherer Ergebnisse (siehe Walker et al., 2002, 2003) mit dem SRTT-Paradigma erfolgreich gelingt. In einer weiteren Untersuchung (Schmidt et al., 2010) wurde die gleiche Aufgabe auf ein an der Wand hängendes Smartboard übertragen. Hier wurde die Sequenz in zwei Experimenten entweder mit nur einer oder mit beiden Händen unter geringen Präzisionsanforderungen ausgeführt. Unter dieser Bedingung fanden sich allerdings keine schlafassoziierten Leistungssteigerungen; ebenso wenig wie bei der DTT, der Kraft-Parametrisierungsaufgabe (CMJ) und der visuomotorischen Verfolgungsaufgabe (für einen Überblick siehe hierzu Blischke et al., 2008).

Aus den Ergebnissen der genannten Untersuchungen (Blischke et al., 2008) geht hervor, dass nicht, wie von Walker et al. (2002) ursprünglich angenommen, alle sensomotorischen Aufgaben schlafassoziierte Konsolidierungsprozesse auslösen, die zu Leistungssteigerungen führen. Vielmehr deuten die Ergebnisse darauf hin, dass die

schlafbasierte Konsolidierung einer neuen Aufgabenrepräsentation sich nur unter bestimmten Bedingungen leistungssteigernd auswirkt. In der Retrospektive kommen auch King et al. (2017) zehn Jahre später in ihrer Übersicht zu einer differenzierten Betrachtung.

2.6 Weitere Untersuchungen alltagsnaher Aufgaben

In der jüngeren Vergangenheit sind Studien weiterer Arbeitsgruppen erschienen, die ebenso Paradigmen mit alltagsnahen Aufgaben zur Anwendung bringen und die vorhandenen Erkenntnisse um einige Aspekte erweitern.

So untersuchten Kempler & Richmond (2012) eine Armpositionierungssequenz unter Einbeziehung beider Arme. Hier galt es für die Probanden, eine aus sechs Einzelementen bestehende Sequenz innerhalb eines 30 Sekunden langen Intervalls so oft wie möglich korrekt auszuführen. Die Koordination der Endeffektoren auf beiden Körperseiten sowie die Anzahl der Entscheidungsalternativen (neun pro Element) dürften die Komplexität der Aufgabe im Vergleich zu klassischen Fingerbewegungssequenzen wesentlich erhöht haben.

Obwohl das experimentelle Design von Kempler & Richmond (2002) in einigen Punkten (bspw. der Anzahl von Aneignungs- und Testblocks) von den klassischen Untersuchungen abweicht, konnte die Studie erstmals zeigen, dass bei einer komplexen großmotorischen Sequenzlernaufgabe Leistungssteigerungen nach Schlaf auftreten.

Al-Sharman & Siengsukon (2013, 2014) nutzten die Verbindung einer Gangaufgabe auf einem für die Probanden neuen ca. 30 Meter langen, in der konkreten Streckenführung unregelmäßig gestalteten, elliptischen Pfad (gefordert war hier also eine fixierte Sequenz spezifischer Richtungsänderungen) – inklusive einer ca. 3 Meter langen Strecke im Tandemschritt – mit einer kognitiven Zweitaufgabe (Rückwärtszählen in 7er-Schritten von einer randomisiert ausgewählten Zahl zwischen 299 und 293), um erstmals eine funktionale motorische Aufgabe auf schlafbasierte Offline-Learning-Effekte zu untersuchen.

Zudem wurden zwei unterschiedliche Populationen untersucht:

1. Junge Erwachsene: ca. 25 Jahre alt (Al-Sharman & Siengsukon, 2013)

2. „Mittelalte“ Erwachsene: ca. 48 Jahre alt (Al-Sharman & Siengsukon, 2014)
3. „Alte“ Erwachsene: ca. 70 Jahre alt (Al-Sharman & Siengsukon, 2014)

Die Autoren fanden, dass sich bei jungen Erwachsenen nur in der Schlafgruppe die Ganggeschwindigkeit, die Schrittlänge im Tandemschritt sowie die Geschwindigkeit des Tandemschritts erhöhten (siehe Al-Sharman & Siengsukon, 2013). Ebenso verhielt es sich für die „mittelalte“ Population (Al-Sharman & Siengsukon, 2014) sowie für die „alten“ Erwachsenen (Al-Sharman & Siengsukon, 2014).

Genzel et al. (2012) führten eine Untersuchung einer sequenziellen Bewegungsaufgabe durch. Die auszuführende Sequenz ähnelte dabei der Abfolge von einfachen Tanzschritten und beinhaltete damit die sequenzielle Koordination der unteren Extremitäten als Reaktion auf visuelle Stimuli.

Das Ergebnis: ein signifikanter Unterschied in der Kriteriumsleistung zwischen beiden Experimentalgruppen nach Schlaf. Diejenige Gruppe, die geschlafen hatte, während die andere eine Wachphase absolvierte, zeigte signifikant bessere Ergebnisse.

Die weitere Untersuchung befasste sich mit der Frage, ob eine erhöhte Leistung nach Schlaf auch bei einem Transfer auf eine neue, bis dato unbekannte Sequenz möglich ist. Hier zeigte sich kein signifikanter Effekt.

Das von Genzel et al. (2012) genutzte Paradigma kann als großmotorische, sequenzielle Reaktionszeitaufgabe mit relativ geringen räumlichen, dafür aber hohen zeitlichen Präzisionsanforderungen charakterisiert werden.

Hödlmoser et al. (2014) untersuchten eine komplexe motorische Adaptationsaufgabe – „die Inverse-Bicycle-Aufgabe“ – auf Leistungssteigerungen nach einem zweistündigen Retentionsintervall, in welchem eine Gruppe schlief, während die andere wach blieb. Im Gegensatz zu den meisten anderen genannten Studien kam hier Tagschlaf zur Anwendung.

Die Aufgabe beinhaltete das Fahren und Steuern eines Fahrrades mit inverser Lenkung. Gemessen wurden sowohl die Fahrgeschwindigkeit als auch die Genauigkeit, mit der die Probanden das Fahrrad durch einen Slalomparcours steuern konnten. Tests wurden unter zwei Bedingungen durchgeführt: einmal beim Geradeausfahren und einmal beim Slalomfahren.

Es fanden sich weder in der Schlaf- noch in der Wachgruppe signifikante Leistungssteigerungen. Allerdings verschlechterte sich die Performance beim Geradeausfahren jeweils nach Wachheit und Schlaf, die Genauigkeit im Slalom jedoch nur nach Schlaf und nicht nach Wachheit.

Morita et al. (2012) untersuchten die Auswirkungen eines 2-stündigen Tagschlafes (Nap) auf Offline-Leistungssteigerungen. Hier wurden zusätzlich Schlafparameter per EEG erhoben. Die verwendete Lernaufgabe war das Jonglieren mit drei Bällen – eine Aufgabe, die aufgrund ihrer Anwendungsnähe und Komplexität gewählt wurde. Das Leistungskriterium war dabei die Anzahl gefangener Bälle in der dem Drei-Ball-Jonglieren unterliegenden Reihenfolge innerhalb von drei Minuten. Die Autoren fanden schlafbasierte Offline-Leistungssteigerungen in der Nap-Gruppe. Veränderungen im EEG (erhöhte Slow Oscillations) gegenüber einem Kontroll-Nap fanden nur in den tiefen Schlafphasen NREM-3 und -4 (Slow-Wave-Sleep – SWS) statt. Zeitbasierte Offline-Leistungssteigerungen in der wachen Gruppe fanden sich dagegen nicht.

Die oben genannten Untersuchungen zeigen, dass schlafassoziierte Leistungssteigerungen auch bei Paradigmen mit anderen Aufgabenanforderungen auftreten als dem ursprünglich genutzten SRTT-Paradigma. Zudem wird deutlich, dass schlafbasierte Leistungssteigerungen auch bei alltagsnäheren, großmotorischen Aufgaben auftreten können. Allerdings konnten Blischke et al. (2008) und Hödlmoser et al. (2014) zeigen, dass sich die Befunde nicht auf beliebige Aufgaben generalisieren lassen.

Die vorliegende Arbeit verfolgt den Ansatz, systematisch zu untersuchen, unter welchen Bedingungen die gefundenen schlafbasierten Leistungssteigerungen auftreten, und wirft einen genaueren Blick darauf, welche Mechanismen und Randbedingungen für diese Effekte verantwortlich bzw. erforderlich sind.

2.7 Begriffsklärung: Verwendung von Terminologien in dieser Arbeit

2.7.1 Großmotorisch

Bezüglich des Begriffes „großmotorisch“ findet sich in der Literatur keine einheitliche Definition. Vielmehr wird der Begriff sehr vielfältig genutzt. „Großmotorische Fertigkeiten“ werden bspw. laut Herrmann et al. (2015) mit Objektmanipulation, Lokomotion oder Gleichgewicht in Verbindung gebracht und sind Gegenstand vieler Untersuchungen zur motorischen Entwicklung von Kindern. Hier wird in Verbindung mit den Begriffen „großmotorische Funktion“ oder „großmotorische Performance“ eine ganze Reihe an unterschiedlichen Bewegungsformen genannt, z.B. werden im „Gross Motor Function Measure“ (Russel et al., 1989) zur Erfassung großmotorischer Fertigkeiten bei Kindern mit cerebraler Lähmung Fertigkeiten wie Gehen, Stehen, Rollen, Sitzen oder Krabbeln genannt (Boyce et al., 1995). Christova et al. (2018) nennen Tätigkeiten wie Tanzen, Golfen oder die Bedienung von Musikinstrumenten als Beispiele für großmotorische Fertigkeiten. Dabei sind laut Christova et al. (2018) größere Körpersegmente beteiligt, die komplexe Muskelsynergien erfordern.

Aufgrund dieser breiten Verwendung des Begriffes scheint es angebracht, ihn für die Benutzung im Rahmen der vorliegenden Arbeit zu konkretisieren.

Der Begriff „großmotorisch“ wird vor hier allem verwendet, um die Abgrenzung zu Aufgabentypen, denen andere motorische Kontrollmechanismen zugrunde liegen, zu verdeutlichen.

Zudem soll die Nutzung des Begriffes „großmotorisch“ darauf hinweisen, dass die motorischen Kontrollanforderungen verschiedener Aufgabentypen sich stark voneinander unterscheiden.

Der Begriff „großmotorisch“ wird in dieser Arbeit daher für diskrete, ungeführte (freie Bewegungen ohne mechanische Einschränkungen) Zielbewegungen verwendet, bei denen mehrere Gelenke involviert sind, von denen wiederum mindestens eines ein proximales Gelenk (d.h. Schulter- oder Hüftgelenk) ist. Dies ist darauf zurückzuführen, dass die Innervationsbedingungen größer, die proximalen Gelenke steuernder Muskelgruppen sich grundsätzlich von denen der Fingermuskulatur unterscheiden und

bspw. das proximale Schultergelenk bei Zielbewegungen mit der Hand immer zuerst aktiv ist. Somit wird auf Ebene der motorischen Kontrolle gewährleistet, dass unter Einbeziehung großer Muskelgruppen mehrere Freiheitsgrade kontrolliert werden müssen und aus der unendlichen Zahl möglicher Bewegungsalternativen häufig diejenige mit den geringsten Kosten ² (z.B. energetischer Art, Kontroll- bzw. Entscheidungsaufwand, potenzielle Korrekturerfordernisse) ausgewählt wird. Es kann allerdings auch eine Bewegungsalternative gewählt werden, die an den jeweiligen Aufgabenbedingungen ausgerichtet und kostenträchtiger als eigentlich notwendig ist.

Diese Definition gewährleistet, dass hinsichtlich der motorischen Kontrollanforderungen ein Unterschied zur Kontrolle von Fingerbewegungssequenzen hergestellt wird. Gleichzeitig findet eine Abgrenzung zur motorischen Kontrolle von rhythmischen Bewegungen, wie bspw. Gangbewegungen, statt, die teilweise durch in der Bewegungskontrollhierarchie niedrigere Instanzen gesteuert werden. Diese sind in der Regel allerdings schon über viele Jahre in der Ontogenese angeeignet und kommen daher nicht als neu zu erlernende Fertigkeiten für den Zweck dieser Arbeit infrage.

2.7.2 Komplexität

Alle oben genannten Untersuchungen haben sich zum Ziel gesetzt herauszufinden, ob schlafassoziierte Leistungssteigerungen auch bei komplexen, alltagsnäheren Lernaufgaben zu finden sind. Eine präzise und einheitliche Definition dessen, welche Elemente einer Aufgabe deren Komplexität determinieren, findet sich allerdings nicht. Auch weitere Literaturstellen geben wenig Aufschluss darüber, was genau eine komplexe Aufgabe kennzeichnet. Laut Wulf & Shea (2002) ist eine allgemeingültige Definition von Aufgabenkomplexität „*a difficult, if not impossible challenge*“ (Wulf & Shea, 2002, p. 185). Es werden zwar diverse Faktoren wie die Bewegungsdauer in Relation zur Distanz und Größe des zu treffenden Zieles, die Anzahl involvierter Freiheitsgrade oder die Anzahl von Sequenzelementen mit dem Begriff „Aufgabenkomplexität“ in Verbindung gebracht, allerdings ist bislang nicht klar, wie verschiedene Aspekte miteinander in Verbindung stehen. Wulf & Shea (2002) gingen daher in ihrem Beitrag dazu über, eine Aufgabe als

² Manchmal entscheidet sich das System auch für eine Variante, die kostenträchtiger ist als notwendig. Bspw. werden häufig gerade Trajektorien bevorzugt. Diese Trajektorien sind allerdings nicht notwendigerweise diejenigen mit den geringsten Kosten.

komplex zu bezeichnen, wenn sie nicht in einer einzigen Aneignungssitzung beherrscht werden kann sowie mehrere Freiheitsgrade beinhaltet. Dahingegen wurden Aufgaben als „simple“ – einfach – eingestuft, die eher als „künstlich“ zu bezeichnen sind (Laboraufgaben im Gegensatz zu anwendungsnahen Aufgaben), nur einen Freiheitsgrad beinhalten und innerhalb einer einzigen Session beherrscht/erlernt werden können. Kuriyama et al. (2004) nutzten in ihrer Untersuchung zum Einfluss der Aufgabenschwierigkeit auf das Ausmaß an Offline-Leistungssteigerungen dagegen eine andere Herangehensweise. Hier wurde der Begriff Komplexität anhand der Anzahl verwendeter Endeffektoren (in diesem Fall Finger), der Anzahl von Auswahlmöglichkeiten (Tastenzahl) sowie der Uni- oder Bilateralität der verwendeten Endeffektoren festgemacht.

Um ein genaueres Verständnis des Zusammenhanges zwischen Aufgabenkomplexität und der Konsolidierung von motorischen Gedächtnisinhalten zu erlangen, erscheint es daher angebracht, den Komplexitätsbegriff weiter zu differenzieren und aus der Perspektive der Kontrollanforderungen an das neuromotorische System in frühen Lernphasen zu betrachten.

In frühen Lernphasen sind im hierarchisch organisierten Bewegungskontrollsystem verstärkt höhere Ebenen – verantwortlich für den Prozess der Identifikation und Koordination der passenden Reaktionen („motor selection“, siehe Diedrichsen & Kornysheva, 2015) – in die Produktion des motorischen Outputs eingebunden (Diedrichsen & Kornysheva, 2015). Aufgabe der höheren Instanzen ist die Aktivierung von motorischen Grundformen („motor primitives“, siehe Diedrichsen & Kornysheva, 2015), die letzten Endes Muskelaktivitäten auslösen. Die Auswahlbene muss mehrere Alternativen berücksichtigen, um das am besten passende Set an „motor primitives“ auszuwählen. Dieser Prozess benötigt Zeit und der Zeitaufwand dafür ist abhängig von der Zahl und Ähnlichkeit der möglichen Antwortalternativen (Diedrichsen & Kornysheva, 2015). Mit zunehmender Übung und Bekanntheit geeigneter „motor primitives“ für die Ausführung einer Aufgabe sinkt die Beteiligung der höheren Instanzen der Kontrollhierarchie an der Bewegungsproduktion und die benötigte Zeitdauer sinkt.

In Bezug auf diesen Mechanismus des motorischen Lernens scheint es sinnvoll, den Begriff der Aufgabenkomplexität in Abhängigkeit von der Anzahl und Ähnlichkeit der zur

Verfügung stehenden Antwortalternativen bei der Ausführung einer Aufgabe – und damit vom Entscheidungsaufwand, den das kognitive Kontrollsystem zu bewältigen hat – zu betrachten. Dies bezieht sich vor allem auf die Aneignung von zu lernenden Bewegungsfolgen (im Unterschied zu den Adaptationsaufgaben, bei denen es darum geht, bereits gelernte Bewegungen an neue Umstände anzupassen), die einen hohen Neuigkeitsgrad für das System haben, sowie künstliche Bewegungsformen.

Unter dieser Prämisse kommen verschiedene Faktoren in Betracht, welche die Komplexität einer fixen Bewegungsfolge (also einer Reihung von Elementen in der immer gleichen Abfolge) determinieren und somit Gegenstand weiterer Untersuchung sein müssen. Bspw.:

- Die Anzahl von Sequenzelementen: Je mehr Elemente, desto höher ist die Anforderung an die höheren Instanzen der internen Kontrollhierarchie.
- Die Sequenzstruktur: Fixe Sequenzen können aus keiner erkennbaren Regel folgenden aneinandergereihten Elementen bestehen. Eine weitere Möglichkeit ist eine Sequenzstruktur, welche die Identifikation von Teilsequenzen erleichtert, indem diese an bekannte Muster und Strukturen bzw. Schemata angelehnt sind. Dies können bei räumlich definierten Zielbewegungsfolgen z.B. geometrische Figuren (regelmäßige Vielecke) oder andere schon gut bekannte räumliche Konfigurationen sein. Durch die Integration dieser bekannten Schemata in die kognitive Repräsentation der abstrakt-räumlichen Bewegungssequenz sinkt der Kontrollaufwand, da bereits interne Repräsentationen der Teilsequenzen vorhanden sein könnten bzw. nicht jedes einzelne Sequenzelement für sich abgerufen werden muss. So kann auch eine Sequenz mit einer vergleichsweise hohen Anzahl an Elementen eine geringe Komplexität aufweisen.
- Die Identifikation und Auswahl von Bewegungszielpunkten: Je höher der visuelle Suchaufwand für ein Bewegungsziel und je größer die Anzahl an Antwortalternativen, desto komplexer ist die Aufgabe.
- Die Anzahl und Koordination der involvierten Endeffektoren, die unabhängig voneinander aktiviert werden: Je mehr Endeffektoren, desto höher ist die Komplexität.

3 Vorgehensweise für die Publikationen der vorliegenden Arbeit

Die bis dato uneinheitliche Datenlage zu schlafassoziierten Konsolidierungseffekten und daraus resultierenden Leistungssteigerungen bei großmotorischen Aufgaben hat den Anlass zur Vorgehensweise für diese Arbeit gegeben. Die Sichtung der Literatur zu schlafbasierten Leistungssteigerungen bei motorischen Aufgaben führt zu folgendem Eindruck:

Offline-Learning scheint besonders ausgeprägt bei Aufgaben aufzutreten, die:

- sequenziell strukturiert sind
- zumindest teilweise deklarative Gedächtniskomponenten beinhalten
- unter expliziten Bedingungen gelernt werden
- eine hinreichende Komplexität besitzen

Umgekehrt zeigen Ergebnisse von Hödlmoser et al. (2014), Blischke et al. (2008), Schmidt (2010), dass Aufgaben, die

- vorrangig zeitliche Präzisionsanforderungen beinhalten,
- unter impliziten Bedingungen angeeignet werden und
- vorrangig Adaptationsanforderungen beinhalten,

keine schlafbasierten Leistungssteigerungen erwarten lassen. Diese Ergebnisse lassen darauf schließen, dass die speziellen Aufgabencharakteristika und damit die Anforderungen an das motorische Kontrollsystem sowie die Art der internen Repräsentation verschiedener Aufgabentypen Faktoren sind, die schlafbasierte Leistungssteigerungen durch die Konsolidierung der internen Repräsentation ermöglichen. Insbesondere scheinen Aufgaben, die vorrangig durch Adaptationsanforderungen gekennzeichnet sind, weniger schlafbegleitenden Optimierungsprozessen (die sich auch in beobachtbaren Leistungssteigerungen ausdrücken) zugänglich zu sein. Bei Adaptationsaufgaben zeigt sich dazu ein gemischtes Bild aus Untersuchungen, die nur teilweise schlafbegleitende Leistungssteigerungen feststellen konnten (siehe hierzu King et al., 2017). Im

Gegensatz dazu finden sich bei Aufgaben, welche das Erlernen räumlich definierter Bewegungsfolgen unter expliziten Bedingungen beinhalten, konsistent schlafbegleitende Leistungssteigerungen (King et al., 2017).

Für das vorliegende Dissertationsvorhaben wurde folgende Vorgehensweise verfolgt:

1. Schritt: Nachweis von schlafbasierten Leistungssteigerungen nach dem Neulernen einer komplexen Armbewegungssequenz, die als großmotorisch bezeichnet wird (Publikation 1)
2. Schritt: Prüfung, welche auf Verhaltensebene nachweisbaren Optimierungen an der Gedächtnisrepräsentation zu den beobachteten Ergebnissen führen (Publikationen 2 & 3)
3. Schritt: Prüfung, welche speziellen Aufgabencharakteristika das Auftreten schlafbasierter Konsolidierungseffekte determinieren (Publikation 4)

Schlafbasierte Leistungssteigerungen bei einer großmotorischen Armbewegungssequenz



Abbildung 1: Grafische Darstellung der dissertationsrelevanten Publikationen.

In den folgenden Unterkapiteln finden sich kurze Zusammenfassungen der jeweiligen Publikation.

3.1 Paradigma, Messapparat, Versuchsaufbau und experimentelles Protokoll für die vorliegenden Publikationen

3.1.1 Das Paradigma

Schmidt et al. (2010) übertrugen die für schlafbasierte Leistungssteigerungen häufig genutzte SRTT in eine großmotorische Variante: Hier musste die Sequenz aus fünf Tastendrücken mit einem Arm (und in einer Variante mit beiden Armen) auf einem vertikal angebrachten Smartboard ausgeführt werden. Allerdings konnten bei dieser Aufgabe keine schlafbasierten Verbesserungen der Kriteriumsleistung beobachtet werden.

Malangré et al. (2014) mutmaßen, dass nicht der großmotorische Charakter der verwendeten Aufgabe für das Ausbleiben der Effekte verantwortlich ist. Vielmehr wird vermutet, dass die geringe Aufgabenschwierigkeit (niedriger Fitts' ID) auf der einen Seite sowie die Möglichkeit der muskulären Ermüdung während der kontinuierlichen Ausführung der Armbewegungssequenz in der Vertikalen auf der anderen Seite dazu führen, dass hier keine schlafbasierten Leistungssteigerungen beobachtbar sind.

Das Peg-Plugging-Paradigma wurde für die vorliegenden Arbeiten gewählt, da es die Ausführung komplexer Armbewegungssequenzen ermöglicht und gleichzeitig die Eliminierung der oben genannten limitierenden Aspekte der von Schmidt et al. (2010) verwendeten Aufgabe zulässt.

Dies erfolgt durch die Verlagerung der Zielpunkte in die horizontale Ebene sowie durch eine eher diskrete Sequenzproduktion und die vergleichsweise hohe Komplexität und Aufgabenschwierigkeit.

Bei dem Peg-Plugging-Paradigma handelt es sich um die Implementierung einer großmotorischen, komplexen Armbewegungssequenz mit hohen Präzisionsanforderungen.

3.1.2 Ablauf der Experimente

Die Probanden sind aufgefordert, nach einem Startsignal in einer festgelegten, sich nicht verändernden Sequenz (je nach Experiment fünf bis 15 Sequenzelemente) mit einem kleinen hölzernen Stift (gehalten im Präzisionsgriff mit der nicht dominanten Hand) verschiedene Endpunkte auf horizontaler Ebene anzusteuern. Die Sequenz soll dabei so schnell und akkurat wie möglich ausgeführt werden. Nach jeder Sequenzausführung erfolgt eine 5-sekündige Pause. Der Start der darauffolgenden Sequenzausführung erfolgt erst auf ein Signal des Versuchsleiters.

Die Peg-Plugging-Aufgabe weist somit bezüglich der Anforderungen an die Bewegungsplanung eher Ähnlichkeiten mit der Discrete-Sequence-Production (DSP) Task (diskrete Sequenz-Produktions-Aufgabe, ursprünglich eine Tastendruckaufgabe zur Ermittlung von Reaktionszeiten; siehe hierzu Verwey et al., 2015) auf als mit der SRTT: Die Sequenzausführung erfolgt jeweils diskret. Somit besteht die Möglichkeit der Vorausplanung der ersten Sequenzelemente vor dem Start der jeweiligen Aufgabenausführung. Durch die diskrete Sequenzproduktion werden zudem kürzere Gesamtbewegungsdauern (zwischen fünf und 18 Sekunden, siehe Malangré et al., 2014) zwischen zwei Ruhephasen erreicht als bei der von Schmidt et al. (2010) verwendeten Aufgabe. Hier lag die Bewegungsdauer zwischen zwei Ruhephasen bei 30 Sekunden. Zusammen mit der Aufgabenausführung in der Horizontalen wurde so versucht, einer etwaigen muskulären Ermüdung vorzubeugen.

3.1.3 Versuchsaufbau

Die Probanden sitzen während der Ausführung der Aufgabe auf einem Stuhl vor einem Tisch, auf dem das elektronische Pegboard aufgebaut ist (siehe Abbildung 2, Abmessungen des Pegboards: Abbildung 3). Dabei ist die sich bewegende Hand permanent sichtbar, während die ruhende Hand neben dem Pegboard auf dem Tisch abgelegt wird. Der Rücken der Versuchspersonen ruht an der Rückenlehne des Stuhles und die Teilnehmer können die involvierten Gelenke (Schulter, Ellbogen und Handgelenk) frei bewegen. Um alle Bewegungen einer Sequenz erfolgreich zu bewältigen, ist es zwingend erforderlich, dass alle drei Gelenke eingesetzt werden.

Der Transport des Stiftes zu den verschiedenen Endpunkten erfordert dabei Armbewegungen über verschiedene Distanzen (3,83 bis 33,75 cm) und in verschiedene

Richtungen (links, rechts, vom Körper weg, zum Körper hin, diagonal) inklusive Bewegungen über die Körpermittelachse hinweg. Die Präzisionsanforderungen für die Teilbewegungen können mit einem Fitts' ID von durchschnittlich 4,95 (Publikation 1) bis 5,35 (Publikation 4) als hoch eingestuft werden. Der Fitts' Index of Difficulty beschreibt die Schwierigkeit von Zielbewegungen anhand der Bewegungsdistanz vom Mittelpunkt der Ausgangslokation bis zum Mittelpunkt der Ziellokation(**A**) sowie der Breite des Bewegungszieles (**W**) anhand folgender Gleichung: $\text{Log}_2(2A/W)$.

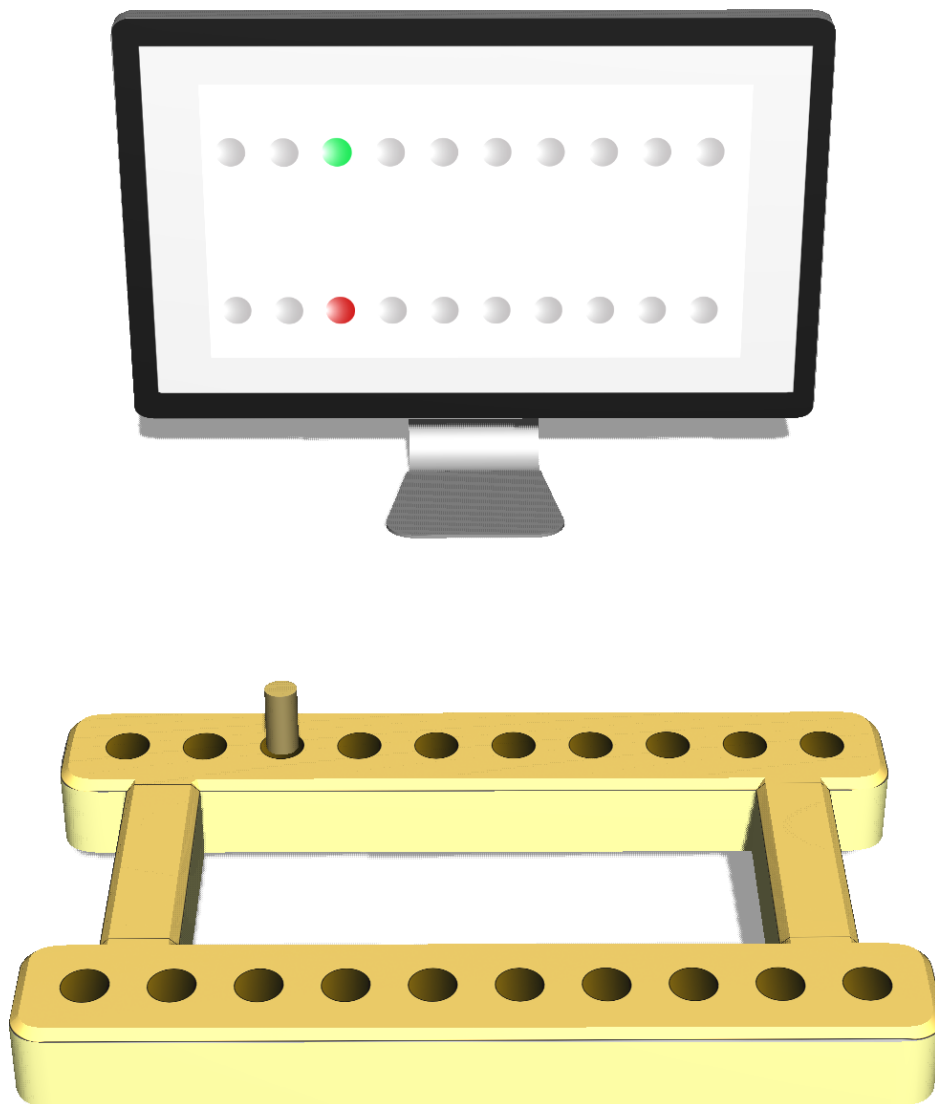


Abbildung 2: Hölzernes Pegboard mit 20 Löchern in zwei Reihen. Auf dem Bildschirm sind die 20 Löcher des Pegboards visuell repräsentiert. Dabei wird die aktuelle Position des Stiftes durch einen

grün gefüllten Punkt markiert. Das nächste zu erreichende Loch wird durch einen roten Punkt repräsentiert.

Der Stift muss am Endpunkt jeder Bewegung in eines der 20 Löcher des Pegboards gesteckt werden, um damit einen magnetischen Kontakt zu schließen.

Der Zielpunkt jeder aktuellen Bewegung wird dabei durch das Aufleuchten eines rot ausgefüllten Kreises auf einem virtuellen Pegboard, welches auf einem Computermonitor abgebildet ist, vorgegeben. Nach dem Schließen eines Kontaktes wechselt die Farbe des Zielpunktes zu Grün und das nächste anzusteuernde Loch wird rot hervorgehoben. Nach Vollendung einer kompletten Sequenz erfolgt eine kurze Pause (ca. 5 s) und nach Äußerung eines „Bereit“-Signales der Versuchsperson folgt eine Wiederholung der gleichen Sequenz.

Die Sequenz wird den Probanden dabei niemals komplett präsentiert, sondern stets lediglich in der beschriebenen Weise. Den Probanden wird allerdings vorab mitgeteilt, dass sie immer die gleiche Armbewegungssequenz durchführen. Während der Aktivität wird den Probanden keinerlei zusätzliche direkte Information über ihre Performance mitgeteilt (keine Mitteilung über Fehler oder Zeiten). Allerdings ist das Erkennen von Fehlern über den Soll-Ist-Abgleich auf dem Bildschirm möglich. Die Abmessungen des Pegboards können aus Abbildung 3 entnommen werden. Für eine detaillierte Beschreibung siehe Malangré et al. (2014).

Aneignungs- und Testphasen sind in Blöcken aus jeweils 10 Sequenzausführungen organisiert. Zwischen jedem Block liegt dabei eine Pause von 30 Sekunden. Um Ermüdungsaufbau vorzubeugen, wurde nach Block 6 eine Pause von 150 Sekunden eingestreut.

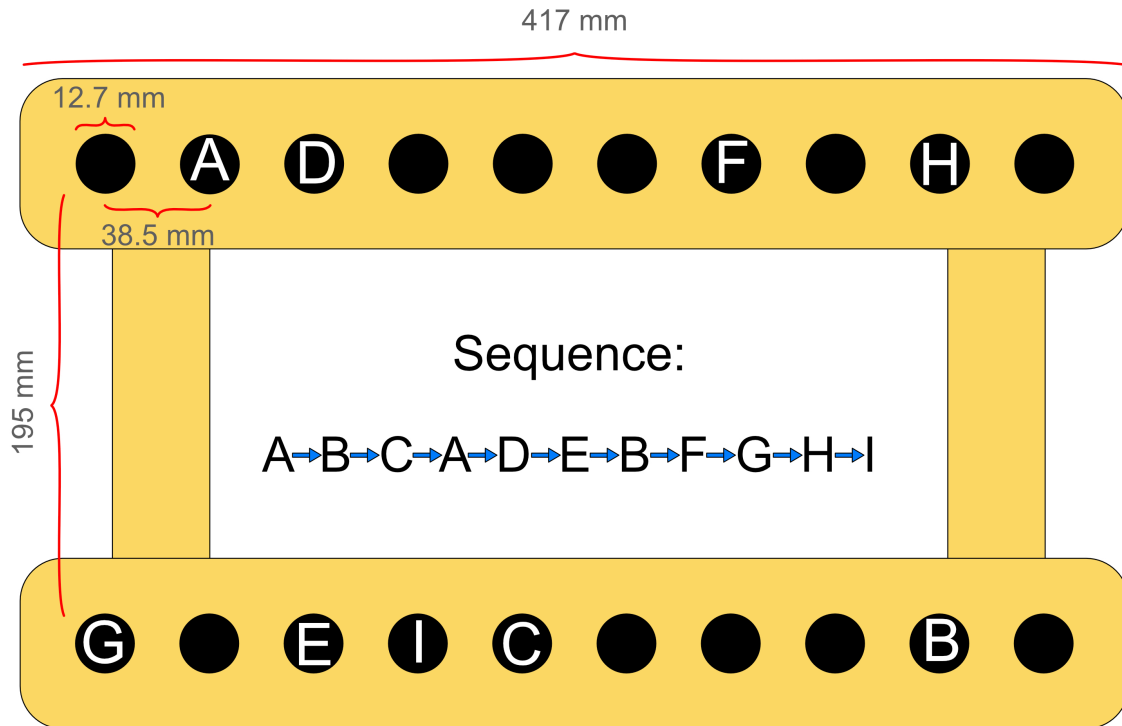


Abbildung 3: Abmessungen des Pegboards und seiner Elemente.

3.1.4 Experimentelles Protokoll

Um nachzuweisen, dass etwaige Leistungssteigerungen tatsächlich schlafassoziiert sind und nicht ebenfalls nach einem Wachintervall auftreten, ist es zwingend notwendig, beide Zeiträume in die Untersuchung miteinzubeziehen. Um sicherzustellen, dass eventuell gefundene Leistungsveränderungen nicht allein auf dem Verstreichen von Zeit beruhen – also unabhängig von Schlaf- oder Wachbedingung in den ersten 12 oder 24 Stunden nach dem Erlernen einer neuen Aufgabe auftreten –, müssen Tages- und Nachtzeiträume in einem Zwei-Gruppen-Versuchsplan balanciert werden. Somit können auch Time-of-Day-Effekte – also unterschiedliche Leistungen aufgrund der Tageszeit zum Zeitpunkt eines Tests – durch den Vergleich der Leistungen am Ende der Aneignungsphase festgestellt werden.

Die Probanden durchliefen das Experimentalprotokoll daher in zwei Gruppen aufgeteilt: die Morning-Evening-Morning(MEM)-Gruppe und die Evening-Morning-Evening(EME)-Gruppe. Die Teilnehmer der Gruppe MEM durchliefen nach einer kurzen Einweisung und einer Gewöhnungsphase an das Pegboard die Aneignungsphase gegen 8 Uhr morgens.

Die Gruppe EME durchlief dasselbe Prozedere, startete allerdings 12 Stunden versetzt, um 20 Uhr abends. Beide Gruppen durchliefen jeweils 12 und 24 Stunden später einen Retest. Die (je nach Experiment) 100 bzw. 120 Wiederholungen in der Aneignungsphase wurden auf Blöcke zu 10 Sequenzausführungen aufgeteilt, zwischen denen jeweils 30 Sekunden Pause gemacht wurde, um etwaigen Fatigue-Effekten (siehe dazu Rickard, 2008) vorzubeugen. Beide Tests enthielten jeweils 30 Wiederholungen der gleichen Sequenz verteilt auf drei Blöcke. Abbildung 4 zeigt eine grafische Darstellung des Versuchsplans.

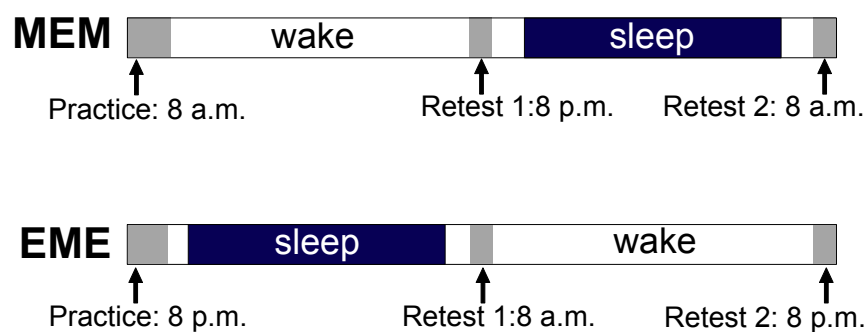


Abbildung 4: Überblick über das Experimentalprotokoll der Untersuchung in Publikation 1. Abkürzungen: **MEM**: „Morning-Evening-Morning“; **EME**: „Evening-Morning-Evening“.

3.1.5 Abhängige Variablen

Die Versuchsanweisung beinhaltete die Aufgabe, eine Sequenz so schnell wie möglich zu absolvieren, allerdings ohne einen Anstieg an Fehlern zu riskieren. Die Geschwindigkeit sollte erst erhöht werden, sobald damit kein Anstieg der Fehlerzahl mehr einherging.

Bei der Datenerhebung wurden primär zwei abhängige Variablen genutzt:

1. ES – *Erroneous Sequences pro Block*: Fehlerhafte Sequenzen pro Versuchsblock à 10 Sequenzausführungen als Fehlermaß. Fehlerhafte Sequenzen wurden nicht wiederholt und es wurden keine Fehler- bzw. Korrekturhinweise gegeben.
2. TET – *Total Execution-Time*: Die Zeitspanne, die für die Ausführung einer kompletten korrekten Bewegungssequenz benötigt wird – vom Verlassen des ersten Loches bis zum Schließen des Kontaktes im letzten Loch. Die TET wurde für jede Versuchsperson über alle korrekten Ausführungsversuche eines

Versuchsblockes gemittelt. Sie ist damit umgekehrt proportional zur mittleren Ausführungsgeschwindigkeit eines Versuchsblockes.

Um sicherzustellen, dass die Leistung der Probanden (TET) während der Aneignungsphase (100 bis 120 Sequenzausführungen) keine Asymptote erreicht, wurde eine Pilotstudie (unveröffentlichte Daten; Schmitz & Waßmuth, 2013) mit 8 Probanden (23,13 \pm 2,1 Jahre; 4 männlich, 4 weiblich) durchgeführt, welche die „Standard“-Peg-Plugging-Sequenz (10 Elemente, siehe hierzu auch Blischke & Malangré, 2017; Supplemental Material, Abbildung S1) zum Gegenstand hatte, welche an drei aufeinanderfolgenden Tagen mit insgesamt 600 Sequenzausführungen (200 pro Tag) absolviert wurde. Das über alle Probanden gemittelte Leistungsmaß (TET) verbesserte sich dabei kontinuierlich bis zur 550. Sequenzausführung. Somit kann davon ausgegangen werden, dass bei der Übungshäufigkeit der in den vorliegenden Publikationen verwendeten Aneignungsphasen (100 bis 120 Sequenzausführungen) noch keine Leistungsasymptote erreicht wird und dass sich das Weiterlernen zu diesem Zeitpunkt immer noch über längere Übungsabschnitte in global beobachtbaren Leistungssteigerungen (TET) niederschlägt.

3.1.6 Zwei Komponenten für die Identifikation von leistungssteigernden Konsolidierungseffekten

Um sicherzustellen, dass Leistungsveränderungen nach einem Behaltensintervall nicht lediglich auf ein etwaiges Weiterlernen durch die Sequenzausführung während der Testzeitpunkte zurückzuführen sind, muss neben dem Vergleich der Leistung zwischen den verschiedenen Testzeitpunkten ein weiterer Vergleich eingeführt werden. Hierfür wurde für jede Versuchsperson über die pro Aneignungsblock gemittelten TET-Daten eine Potenzfunktion gelegt, die den ungefähren Verlauf der jeweiligen individuellen Aneignungskurve beschreibt. Mittels Extrapolation der Potenzfunktion wurde geschätzt, wie sich die Leistung (TET) jeder Versuchsperson entwickelt hätte, wenn die Aneignungsphase sechs Blöcke mehr enthalten hätte (Beispiel: siehe Abbildung 5). Siehe hierzu bspw. auch Pan & Rickard (2015).

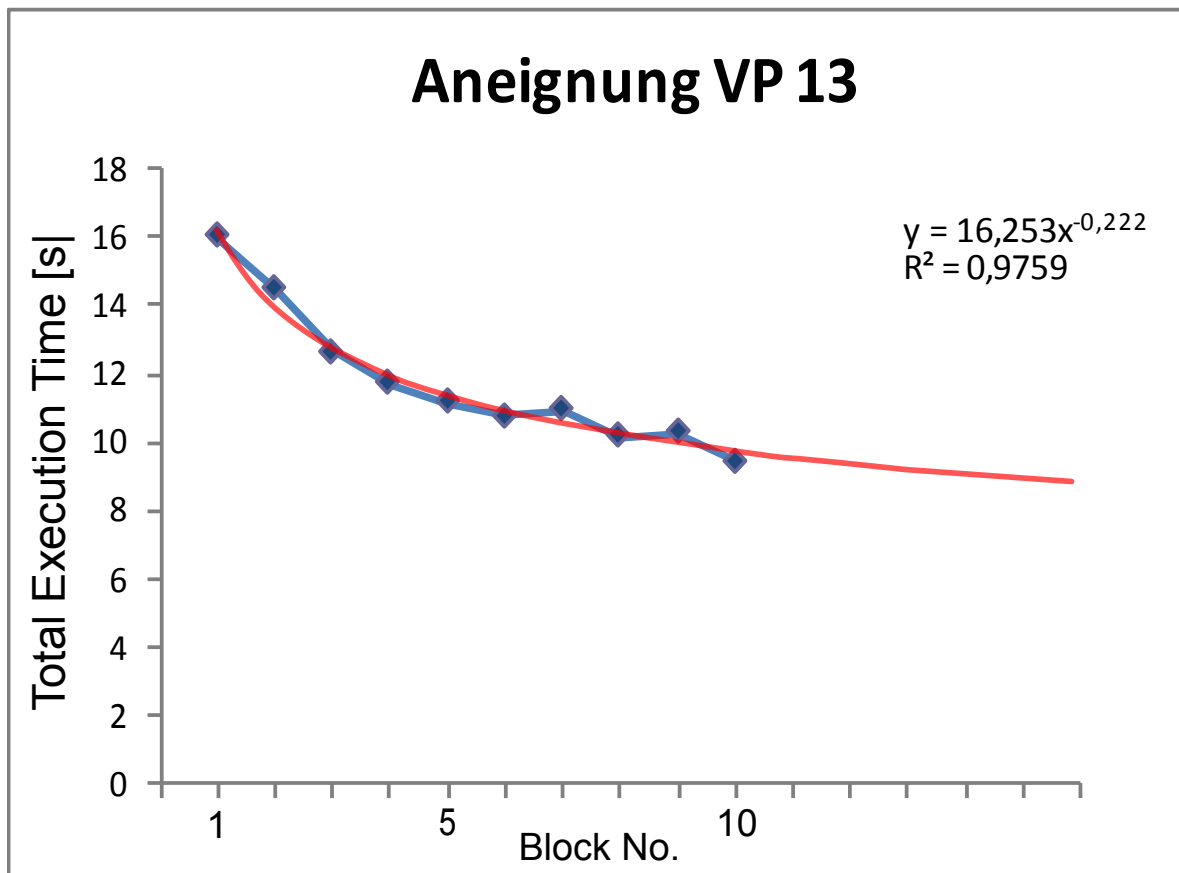


Abbildung 5: Beispiel für den Leistungsverlauf (blockweise gemittelte TET-Werte für die Blöcke 1 bis 10; blaue Linie) in der Aneignungs- und Testphase einer Versuchsperson, inklusive daran angepasster und für die Blöcke 11 bis 16 extrapolierte Potenzfunktion (rote Linie). Durch die abgebildete Formel wird die Potenzfunktion beschrieben.

Somit ergeben sich zwei Komponenten für die Identifikation schlafassoziierte Leistungssteigerungen:

- a) Eine signifikante Verbesserung der Leistung zwischen zwei Messzeitpunkten
- b) Ein signifikanter Unterschied zwischen der tatsächlichen Leistung und der aus dem Aneignungsverlauf extrapolierte Prognose für die weitere Aufgabenausführung in den nächsten Blöcken aus den geschätzten Ausführungsdauern im Retest

Eine schlafbasierte Leistungssteigerung kann nur dann attestiert werden, wenn bei beiden Komponenten signifikante Unterschiede festgestellt werden und anhand der deskriptiven Daten eine Verbesserung der Kriteriumsleistung offensichtlich ist.

4 Überblick über die Publikationen

4.1 Publikation 1: Sleep-related offline learning in a complex arm movement sequence (Malangré, Leinen & Blischke, 2014)

4.1.1 Ziel

Ziel von Publikation 1 war der Nachweis von schlafbasierten Leistungssteigerungen nach dem Neulernen einer großmotorischen, komplexen Armbewegungssequenz und damit einer ein Stück weit näher an alltäglichen Bewegungsaufgaben liegenden, sequenziell strukturierten Bewegungsaufgabe, als es bei Fingersequenzaufgaben der Fall ist. Zudem bildete der in Publikation 1 erbrachte Nachweis die Basis für die in den darauffolgenden Publikationen beschriebenen weiteren Experimente, deren Ziel es war, einen genaueren Blick auf die den schlafbasierten Leistungssteigerungen unterliegenden Mechanismen zu werfen.

4.1.2 Aufbau

Für Publikation 1 wurde das oben beschriebene Peg-Plugging-Paradigma in folgender Konfiguration genutzt: 24 junge Erwachsene (25,5 +/- 3,9 Jahre; 6 weiblich, 18 männlich, eine Linkshänderin) wurden den Gruppen EME (n = 12) und MEM (n = 12) per Pseudorandomisierung zugeteilt. Die fixe Sequenz bestand aus 15 Elementen. Der Abfolge der Sequenzelemente lag keine erkennbare Regelmäßigkeit zugrunde. Die Aneignungsphase bestand aus 100 Sequenzausführungen (10 Blöcken zu jeweils 10 Sequenzen) und zwei Tests: Retest 1 nach jeweils 12 Stunden enthielt 30 Sequenzausführungen aufgeteilt auf drei Blöcke. Retest 2 enthielt 30 weitere Sequenzausführungen aufgeteilt auf drei Blöcke und fand nach weiteren 12 Stunden statt.

4.1.3 Ergebnisse

In beiden Gruppen verkürzte sich die Ausführungsdauer (TET) signifikant (entspricht einer Leistungsverbesserung) während des Aneignungszeitraumes. Darüber hinaus fand eine weitere signifikante Leistungssteigerung (Verringerung der TET) ausschließlich im Test nach dem jeweils durchgeschlafenen Zeitraum statt, nicht jedoch nach dem respektiven Wachintervall. Gleichzeitig war die Leistung nach Schlaf signifikant besser als es anhand der Extrapolation des Aneignungsverlaufes zu erwartenden war. Bezüglich des Fehlermaßes (ES) ergaben sich im Behaltenszeitraum keine signifikanten Veränderungen. Damit wurde in Publikation 1 der Nachweis erbracht, dass auch bei einer großmotorischen Armbewegungssequenz schlafbasierte Leistungssteigerungen auftreten. Die Ergebnisse von Publikation 1 bestätigen, dass auch bei großmotorischen, alltagsnäheren Aufgaben die ursprünglich von Walker et al. (2002) gefundenen schlafbedingten Leistungssteigerungen auftreten können.

4.2 Publikation 2: Chunk Concatenation evolves with practice and sleep-related enhancement consolidation in a complex arm movement sequence (Blischke & Malangré, 2016)

4.2.1 Ziel

Während einige mit schlafspezifischen Leistungssteigerungen in Verbindung stehende neuronale Mechanismen (bspw. Replay, synaptische Homöostase, Schlafspindeln, Veränderung von Schlafparametern (siehe Genzel et al., 2013; Genzel & Robertson, 2015; Born & Wilhelm, 2012)) aktiv beforscht werden, ist noch relativ unklar, welche konkreten verhaltensrelevanten Veränderungen an der spezifischen Gedächtnisrepräsentation einer Bewegungsfertigkeit für die beobachteten Leistungssteigerungen verantwortlich sind. Konkret gilt es, der Frage nachzugehen, welche auf Verhaltensebene zu beobachtenden schlafbasierten Veränderungen sich als konkrete Hinweise auf eine verhaltensrelevante Optimierung an der unterliegenden Gedächtnisrepräsentation einer neu gelernten motorischen Bewegungssequenz interpretieren lassen.

4.2.2 Theoretische und empirische Ausgangslage

Eine Literaturdurchsicht zeigt, dass die Frage nach verhaltensseitig beobachtbaren Veränderungen, die auf konkrete Optimierungen an der Gedächtnisrepräsentation einer spezifischen Bewegungsfertigkeit schließen lassen, bis dato nur selten gestellt wurde. Lediglich Kuriyama et al. (2004) untersuchten schlafbasierte Veränderungen der qualitativen Aufgabenausführung bei unterschiedlich komplexen Variationen einer sequenziellen Fingertapping-Aufgabe.

Kuriyama et al. (2004) konnten zeigen, dass die zeitlichen Abstände zwischen zwei Tastendrücken innerhalb einer Sequenz unterschiedlich lang sind – ein bei der Ausführung von Bewegungssequenzen und anderen Aufgaben wie der Sprachproduktion sehr gut dokumentiertes Phänomen, das in der Literatur vielfach als Indikator für die Segmentierung von Bewegungsfolgen („Chunking“) bezeichnet wird (siehe auch Abrahamse et al., 2013; Verwey et al., 2015; Sakai et al., 2003).

Chunking bezeichnet dabei das Aufteilen einer kompletten Sequenz in mehrere kurze Subsequenzen, die von einem einzigen motorischen Programm vorbereitet und ausgeführt werden können (Blischke & Malangré, 2016; Verwey, 2015).

Innerhalb einer Bewegungssequenz aus verschiedenen Tastenanschlägen ist ein Chunk anhand der unterschiedlich langen Zeitdauern zwischen zwei Tastenanschlägen identifizierbar (im weiteren Verlauf einfach Elementdauer genannt). Ein Chunk wird durch die Kombination eines Elementes mit einer langen Übergangsdauer zwischen zwei Tastenanschlägen, gefolgt von einem oder mehreren aufeinanderfolgenden Elementen mit kurzen Übergangsdauern, charakterisiert. Die lange Dauer für die Ausführung des ersten Elementes reflektiert dabei die Zeitdauer, die – so die zugrundeliegende Modellannahme – der kognitive Prozessor für die Auswahl und das Laden des jeweiligen Chunks in einen motorischen „Puffer“ (für weitere Informationen hierzu siehe Verwey et al., 2015) benötigt. Sie fällt an, da dieser Prozess (zumindest bei noch geringer Expertise bezüglich einer neu zu lernenden Bewegungsfertigkeit) nicht zeitgleich mit der Ausführung vorheriger Sequenzelemente bearbeitet werden kann. Die direkt darauffolgenden Elemente werden als integrierte Teile dieses Chunks betrachtet und benötigen keine weiteren zeitkonsumierenden planerischen Prozesse, sondern reflektieren lediglich die für die motorische Exekution durch den motorischen Prozessor benötigte Zeitdauer (siehe hierzu auch Abrahamse et al., 2013). Da ein Chunk aufgrund

von Limitierungen der Informationsverarbeitungskapazität nur eine finite Zahl an Einzelitems beinhalten kann, werden längere Sequenzen in mehrere Chunks aufgeteilt. Dieser Prozess wird als „*chunk-formation*“ bezeichnet. *chunk-formation* und *Concatenation* (d.h. ein optimierter Übergang zwischen zwei aufeinanderfolgenden Chunks) können also durch die Entlastung der Kapazität für Planung und Ausführung von Sequenzelementen die Performance der gesamten Sequenzausführung positiv beeinflussen.

Kuriyama et al. (2004) identifizierten in ihrer Untersuchung daher sogenannte „problem points“ – die jeweils individuell langsamsten Übergänge von einem Tastenanschlag zum nächsten innerhalb einer Sequenz – sowie die jeweils schnellsten Übergänge („easy points“) und beobachteten deren Entwicklung über den Verlauf von Aneignung und Retests. Speziell von Interesse war dabei, ob sich die Elementdauern der „problem points“ nach einem durchschlafenen Behaltensintervall stärker verändern als die Übergangsdauern der „easy points“. Hier zeigte sich, dass sich die Übergänge an den „problem points“ nur nach einer durchschlafenen Nacht stärker verkürzten als die Übergänge an den „easy points“. Im Gegenzug trat dieses Phänomen bei einem im Wachzustand verbrachten Behaltensintervall nicht auf.

Die Daten von Kuriyama et al. (2004) lassen darauf schließen, dass schlafbasierte Konsolidierungsprozesse dafür verantwortlich sind, dass Leistungssteigerungen über Nacht vor allem durch im Vergleich zu den „easy points“ stärker verkürzten Elementdauern an den „problem points“ zustande kommen.

Diese verringerten Zeitdauern an den „problem points“ deuten auf ein Phänomen hin, das als „*Chunk-Concatenation*“ – ein optimierter Übergang zwischen zwei aufeinanderfolgenden Chunks – bezeichnet wird. Dabei ist aufgrund von Limitierungen des Arbeitsgedächtnisses weniger davon auszugehen, dass alle Informationen zweier aufeinanderfolgender Chunks in einen einzigen Chunk integriert werden. Wahrscheinlicher ist, dass Identifikation, Laden und Initiierung eines Chunks insofern optimiert werden, als der kognitive Prozessor immer mehr von Detailschritten, wie der Spezifikation von Bewegungsparametern (bspw. Bewegungsrichtung, Kraftaufwand, oder Gelenkwinkelstellung), befreit wird (siehe hierzu auch Verwey et al., 2015; Blischke & Malangré, 2016).

Die Ergebnisse von Kuriyama et al. (2004) lassen vermuten, dass Konsolidierungsprozesse während des Schlafes dafür sorgen, dass die Repräsentation einer Aufgabe genau an dieser Stelle optimiert wird und zu den beobachteten Leistungssteigerungen führt.

Publikation 2 beschäftigt sich daher mit der Frage, ob Chunking – in Form von Chunk-Formation und Chunk-Concatenation – auch bei der in Publikation 1 genutzten großmotorischen Armbewegungssequenz am Pegboard auftritt und ob, ebenso wie bei Kuriyama et al. (2004), schlafbasierte Konsolidierungsprozesse Chunk-Concatenation zur Folge haben.

Dies hat insofern Neuigkeitswert, als – zumindest nach dem Kenntnisstand der Autoren – im Kontext der Beforschung schlafassoziierten Konsolidierungsprozesse Chunk-Formation bislang noch nicht bei anderen motorischen Fertigkeiten außer Reaktionszeitdaten dokumentiert wurde.

4.2.3 Methode

Dies wurde in der vorliegenden Arbeit dadurch ermöglicht, dass die Transportbewegungen der Aufgabenelemente (bzw. deren elementspezifische Bewegungsdauern) durch eine Relativierungsprozedur (siehe dazu Blischke & Malangré, 2016) unabhängig vom Einfluss der jeweiligen Trajektorienlänge miteinander vergleichbar gemacht wurden. Hierfür wurde jede der 15 Einzelbewegungen mehrfach von drei Experten isoliert und somit unabhängig vom Sequenzkontext mit maximaler Geschwindigkeit ausgeführt. Anschließend wurden die tatsächlichen gemittelten individuellen Einzelbewegungsdauern (Element-Movement-Durations – EMD) der Probanden von Publikation 1 anhand der mittleren EMTs des jeweiligen Sequenzelementes relativiert. So entstand die dimensionslose Einheit Relativized Element-Movement-Duration (rEMD), welche als Basis für die weiteren Berechnungen diente.

Anders als Kuriyama et al. (2004) wurden hier allerdings statt nur zwei Transitionsunkten (längste vs. kürzeste Reaktionszeit) zwei ganze Elementsets miteinander verglichen. Es wurden insgesamt 4 Elemente mit langen relativierten Elementdauern als Concatenation-Points identifiziert und (für die genaue Prozedur siehe Blischke & Malangré, 2016) mit jeweils einem oder mehreren darauffolgenden „kurzen“

Elementen der Sequenz (welche als integrale Bestandteile eines Chunks gewertet wurden) identifiziert. Es wurden also alle Sequenzelemente in die Analyse miteinbezogen. Damit kann die spezifische Chunk-Struktur auch für die hier verwendete großmotorische Zielbewegungssequenz dokumentiert werden.

4.2.4 Ergebnisse

Für Publikation 2 wurden die Daten aus Publikation 1 zur Re-Analyse auf Einzelelementebene genutzt und gezeigt, dass Chunk-Formation auch bei einer großmotorischen Armbewegungssequenz am Pegboard auftritt. Zudem wurde gezeigt, dass die Veränderungen in der TET (siehe Experiment 1) eindeutig und vollständig auf die Veränderung der Dauern der Concatenation-Points zurückzuführen sind. Somit wurde der bestehende Ansatz zur Kennzeichnung von Chunking-Strukturen und Concatenation-Prozessen erweitert. Daneben wurde die inhaltlich begründete Einteilung von Elementklassen in lange rEMTs vs. kurze rEMTs zum Nachweis von Chunk-Formation noch auf einem weiteren, davon unabhängigen und rein formalen Weg, dem k-Means-Clustering (für eine genaue Beschreibung der Methode siehe Blischke & Malangré, 2016, S. 16), eingeteilt. Die Ergebnisse des k-Means-Clustering bestätigten dabei die Einteilung in Concatenation-Points und darauffolgende Bestandteile eines Chunks.

Entgegen den Ergebnissen anderer Studien (bspw. Kuriyama et al., 2004) fand sich bei der genutzten Peg-Plugging-Sequenz allerdings ein überwiegend nicht idiosynkratisches Muster der langen und kurzen rEMTs. Weitere Publikationen (Sakai et al., 2003; Kuriyama et al., 2004; siehe aber auch Verwey & Eikelboom, 2003³) fanden individuell unterschiedliche „problem points“ bei verschiedenen Probanden. Aufgrund der sich von Fingertapping-Sequenzen deutlich unterscheidenden Peg-Plugging-Aufgabe lässt sich das in Publikation 2 gefundene non-idiosynkratische Muster allerdings anhand der spezifischen Aufgabenbedingungen erklären.

Weiterhin konnte in Publikation 2 nachgewiesen werden, dass beim Peg-Plugging-Paradigma – und damit bei einer großmotorischen, komplexen Bewegungsfolge mit räumlich unterschiedlich definierten, unterschiedlich langen Trajektorien – neben Chunk-Formation auch Chunk-Concatenation auftritt, und zwar sowohl während der

³ Interessanterweise fanden auch Verwey & Eikelboom (2003) unter bestimmten Bedingungen nicht idiosynkratische Sequenzmuster (siehe Verwey & Eikelboom, 2003, p. 178, Figure 3)

Aneignungsphase als auch nach einem Schlafintervall. Chunk-Concatenation trat jedoch nicht nach einer Wachphase auf.

Diese Ergebnisse stehen teilweise in Einklang mit denen von Kuriyama et al. (2004), die ebenso Chunk-Concatenation nach einem Schlafintervall fanden. Allerdings trat Chunk-Concatenation bei der von ihnen genutzten SRTT nach Interpretation der Autoren (Kuriyama et al., 2004) nicht während der Aneignungsphase auf.

Zusammen mit den Ergebnissen von Kuriyama et al. (2004) bilden die Ergebnisse aus Publikation 2 einen eindrücklichen Hinweis darauf, dass während des Schlafes nach dem Neulernen einer motorischen Aufgabe Konsolidierungsprozesse stattfinden, die eine qualitative Optimierung (wahrscheinlich durch optimiertes Identifizieren und Laden von Chunks innerhalb der Sequenz) an der Gedächtnisrepräsentation neu gelernter Aufgaben zur Folge haben, und dass Chunk-Concatenation nach einer durchschlafenen Nacht maßgeblich zu den beobachteten Leistungsgewinnen beiträgt.

4.3 Publikation 3: Sleep-related offline improvements in gross motor task performance occur under free recall requirements (Malangré & Blischke, 2016)

4.3.1 Ziel

Eine weitere offene Frage bezüglich der gefundenen Leistungssteigerungen ist, welche Einflussgrößen über den Verlauf einer durchschlafenen Nacht einer Veränderung unterliegen: die Gedächtnisrepräsentation spezieller Aufgabenelemente (bspw. die Repräsentation der Reihenfolge von Bewegungselementen oder die Repräsentation von Bewegungsvorschriften) oder andere Faktoren, wie bspw. die Reaktions- und Verarbeitungsdauer eines spezifischen visuellen Stimulus oder das Blickverhalten (bspw. bei der Peg-Plugging-Aufgabe das Zielsuchen und die Zielerfassung auf dem Bildschirm; der Blickwechsel von der Informationsquelle hin zum „Arbeitsraum“; die Zielsuche und -erfassung im Arbeitsraum / auf dem Pegboard)? Oder ob sich die gefundenen Effekte ganz einfach auf das verbesserte Weiterlernen in den Retests bei Verfügbarkeit visueller Stimuli und visuellen Feedbacks zurückführen lassen. Einige Veröffentlichungen deuten darauf hin, dass einerseits während des Schlaf-Replays eine aktivitätsunabhängige (Re-)Aktivierung der gleichen Neuronennetzwerke wie bei der tatsächlichen

Aufgabenausführung stattfindet und damit einhergehend eine Verlagerung der Repräsentation in unterschiedliche Gehirnnareale (siehe Genzel et al., 2015; Born & Wilhelm, 2011). Somit liegt die Vermutung nah, dass die verhaltensseitig beobachteten Optimierungen tatsächlich auf eine Veränderung der Gedächtnisrepräsentation der Bewegungssequenz (bzw. der Bewegungsvorschrift, welche die Ausführung der Sequenz leitet und die ohne extrinsische Stimulusinformation verfügbar ist) zurückzuführen sind.

In Publikation 3 wurde daher überprüft, ob die in Publikation 1 gefundenen Leistungssteigerungen ebenso unter Free-Recall-Bedingungen – also unter Entzug jeglicher Stimulusinformation bezüglich der auszuführenden Sequenz während der Behaltenstests – auftreten.

Durch den Entzug der Stimulusinformation wird ausgeschlossen, dass 1.) die Produktion der Sequenz lediglich als eine Reaktion auf visuelle Trigger zustande kommt und 2.) ein informationell gestütztes Weiterlernen während der Retentionstest als Einflussfaktor für die bisherigen Befunde ausgeschlossen werden kann. Vielmehr muss die Sequenz ausschließlich auf der Basis der vorhandenen Gedächtnisrepräsentation der Bewegungssequenz ausgeführt werden. Somit stellt Publikation 3 eine weitere Annäherung an die Anforderungen von Alltagsaufgaben, wie sie in Rehabilitation, Berufsalltag oder Sport vorkommen, dar. In solchen Situationen stehen üblicherweise während der Aneignungsphase Sollwertinformation und Feedback zur Verfügung. Bei der Anwendung des Gelernten sind Instruktion und Feedback allerdings häufig nicht mehr verfügbar, bspw. in therapeutischen Kontexten oder beim Abruf gelernter Bewegungsmuster im sportlichen Wettkampf. Dies ist besonders hervorzuheben, da die vorliegende Arbeit versucht, bisherige Erkenntnisse auch an alltagsnäheren Aufgaben nachzuweisen.

Sollte eine Leistungssteigerung auch unter diesen Umständen auftreten, wäre dies als solider Hinweis darauf zu werten, dass schlafbasierte Leistungssteigerungen in erster Linie auf Optimierungen an der Gedächtnisrepräsentation zurückzuführen sind und nicht auf informationell gestütztes Weiterlernen während der Retentionstests oder auf eine optimierte Erfassung und Anwendung der visuellen Stimulusinformation. Ob die Effekte exklusiv auf Optimierungen an der Repräsentation zurückzuführen sind, bliebe weiter zu klären.

4.3.2 Methode

Für die Untersuchung zu Free Recall wurde die Anzahl der Sequenzelemente auf zehn gesenkt, um die Menge zu encodierender Information zu verringern. Dazu wurde die 15 Elemente umfassende Sequenz aus Publikation 1 um die letzten 5 Elemente gekürzt, so dass eine zehn Elemente umfassende Sequenz entstand. Gleichzeitig wurde die Anzahl der Trials während der Aneignungsphase auf 120 erhöht, um zu gewährleisten, dass tatsächlich eine korrekte Repräsentation der Sequenz innerhalb der Aneignungsphase aufgebaut werden konnte. Diese Erhöhung basierte auf der Erfahrung aus der oben genannten Pilotstudie mit 600 Trials (unveröffentlichte Daten; Schmitz & Waßmuth, 2013). Hier wurde jeweils nach 100 Sequenzausführungen ein Free-Recall-Test durchgeführt, um abschätzen zu können, ab welcher Anzahl von Trials die Probanden die auch für Publikation 3 genutzte 10er-Sequenz verlässlich ohne die Verfügbarkeit visueller Stimuli reproduzieren konnten. Es zeigte sich, dass die VPs nach 100 Sequenzausführungen die Sequenz noch nicht zuverlässig auswendig reproduzieren konnten, nach 200 Sequenzausführungen jedoch schon.

Der eigentlichen Untersuchung vorgeschaltet war eine auch in Publikation 3 erwähnte Pilotstudie. In dieser Pilotstudie wurde aufgrund der Erfahrungen aus der Pilotstudie mit 600 Trials die Anzahl der Trails in der Aneignungsphase auf 120 erhöht, um zu erfahren, ob diese Anzahl an Sequenzausführungen ausreicht, um die Sequenz zuverlässig nach einem Behaltensintervall reproduzieren zu können. Die Untersuchung beinhaltete zwei Gruppen mit einem jeweils nur 12-stündigen Behaltensintervall (einmal über Nacht: Gruppe AM ($n = 9$); einmal über Tag: Gruppe MA ($n = 11$)). Anders als in der Hauptuntersuchung wurde hier nur der Behaltenstest unter Free-Recall-Bedingungen durchgeführt. Als Referenzwerte für die Analyse von Leistungsveränderungen über den Behaltenszeitraum wurden hier lediglich die über die letzten drei von insgesamt 12 informationell gestützten Trial-Blocks á jeweils 10 Trials pro VP gemittelten Ausführungsleistungen herangezogen. Ein Kurzzeit-Behaltenstest war nicht vorgesehen. Ziel war es, zunächst herauszufinden, ob

1. die Sequenz (ohne einen zusätzlichen Kurzzeit-Behaltenstest) nach einem 12-stündigen Behaltensintervall (Schlaf vs. Wachheit) überhaupt korrekt reproduziert werden kann;

2. bei korrekter Wiedergabe unter Free-Recall-Bedingungen eine Leistungssteigerung, -stabilisierung oder -verschlechterung gegenüber dem Ende der Aneignungsphase auftritt.

Die Ergebnisse dieser Pilotstudie offenbarten, dass in der Gruppe „Morning-Evening“ lediglich 2 von 9 Probanden unter Free-Recall-Bedingungen in der Lage waren, die Sequenz korrekt auszuführen. Demgegenüber waren in der „Evening-Morning“-Gruppe neun von elf Probanden dazu in der Lage. In dieser Pilotstudie konnte im Unterschied zu Publikation 1 in der „Evening-Morning“-Bedingung nach 12 Stunden im Free-Recall-Test nur eine Stabilisierung der Leistung festgestellt werden. Demgegenüber waren nach einem entsprechenden Wachintervall massive Vergessenseffekte zu beobachten, die einen quantitativen Leistungsvergleich beider Untersuchungsbedingungen hinsichtlich der Sequenzausführungsgeschwindigkeit unmöglich machten. Offensichtlich waren die 120 Ausführungsversuche während der Aneignungsprozedur nicht ausreichend, um die 10-teilige Armbewegungssequenz ohne schlafbasierte Konsolidierungsprozesse behalten oder gegenüber im weiteren Tagesverlauf auftretenden interferierenden Einflüssen stabilisieren zu können.

Für das weitere Vorgehen boten sich nun zwei Möglichkeiten: Entweder konnte die Anzahl der Aneignungsversuche deutlich gesteigert werden oder durch geeignete Maßnahmen versucht werden, die Verarbeitungstiefe der zu behaltenden Information zu intensivieren; dies allerdings, ohne zu riskieren, dass die VPs durch zu viel informationell gestützte Aneignungsprozedur an eine Leistungsasymptote gelangten.

Für die erste Alternative sprach die Erfahrung aus der oben genannten Pilotstudie mit 600 Trials, dass eine verlässliche Sequenzreproduktion erst ab 200 Trials stattfindet. Derartige Übungsraten lassen sich aber kaum dem Stadium des initialen Bewegungslernens zurechnen, auf das sich die hier gezeigten Studien bezüglich schlafassoziierten Leistungssteigerungen bislang bezogen haben. Für das weitere Vorgehen wurde von den Autoren daher die zweite Alternative gewählt. In Publikation 3 wurde dementsprechend das Experimentalprotokoll leicht verändert. Die Anzahl an Sequenzausführungen in der Aneignungsphase und den Tests blieb unverändert, jedoch wurde 15 Minuten nach der Aneignungsphase, d.h. noch innerhalb der unmittelbar folgenden 4 bis 6 Stunden dauernden Konsolidierungsphase (siehe Brashers-Krug et al., 1996), ein Kurzzeit-Free-Recall-

Test hinzugefügt – drei Blöcke zu zehn Sequenzausführungen ohne Verfügbarkeit visueller Stimuli.

Folgende Überlegungen waren für die Einführung des Kurzzeit-Free-Recall-Tests ausschlaggebend:

Durch die Free-Recall-Bedingung sollte ein erzwungenes aktives Retrieval nach Entleeren des Kurzzeitgedächtnisses (15 Minuten Behaltensintervall bei Verhinderung von aktivem Rehearsal) eine hochintensive Encodierung der zu behaltenden Informationen bewirken, bevor diese durch Interferenzvorgänge gestört werden konnte. Zudem ist aus der Literatur bekannt, dass bei freier Reproduktion von Aufgaben bzw. dem freien Abruf von Informationen der sogenannte „Testing-Effekt“ zum Tragen kommt (siehe bspw. Roediger & Karpicke, 2006; Carpenter & Kelly, 2012; Antnoy & Paller, 2018). Der Testing-Effekt besagt, dass ein Test, also der Abruf encodierter Informationen ohne Feedback und ohne Stimulusinformation, zu einer Verbesserung der Abrufleistung zu späteren Zeitpunkten führt (bspw. am nächsten Tag). Des Weiteren diene der Kurzzeit-Free-Recall-Test als Referenzpunkt für den Vergleich der Leistungen vor und nach dem – entweder durchwachten oder durchgeschlafenen 12-stündigen – Behaltensintervall. Somit wurde gewährleistet, dass dieser Vergleich unter gleichen informationellen Bedingungen durchgeführt wurde und zu keinem der Testzeitpunkte informationell gestütztes Weiterlernen stattfand.

Und schlussendlich diene der Kurzzeit-Free-Recall-Test dazu, sicherzustellen, dass Sequenzproduktion unter der Free-Recall-Bedingung 15 Minuten nach Ende der Aneignungsphase überhaupt möglich war.

4.3.3 Ergebnisse

Die Ergebnisse des Experimentes von Publikation 3 zeigen, dass sich – wie in den anderen Experimenten auch – die Leistung der Probanden bezüglich der Ausführungsdauer (TET) während der Aneignungsphase signifikant verbesserte, während die Fehlerrate (ES) konstant blieb. Zwischen dem Ende der Aneignungsphase und dem Kurzzeit-Free-Recall-Test fand keine signifikante Leistungsveränderung statt (TET und ES). Der Kurzzeit-Behaltenszeitraum von 15

Minuten hatte damit weder Leistungseinbußen zur Folge, noch ließ sich ein Hinweis auf ein „Early Boost“-Phänomen (siehe Ausblick) erkennen. Im Gegensatz dazu zeigten sich weitere Leistungssteigerungen (signifikant reduzierte TET) nur nach einem schlafgefüllten 12-stündigen Behaltensintervall. Nach dem jeweils durchwachten Behaltensintervall blieb die Leistung unverändert.

Publikation 3 konnte damit zeigen, dass schlafbasierte Leistungssteigerungen bei einer großmotorischen Aufgabe auch dann auftreten, wenn während der Aufgabenausführung nach einem Behaltensintervall keine visuellen oder anderen Informationen über die Kriteriumsaufgabe (bspw. Instruktionen und Feedback) zur Verfügung stehen.

Daraus ergeben sich folgende Erkenntnisse:

1. Die korrekte Reproduktion der Sequenz unter der Free-Recall-Bedingung scheint in Anbetracht der hier gewählten – vergleichsweise geringen – Anzahl an Sequenzausführungen in der Aneignungsphase und der Dauer des Behaltensintervalls durch einen „Kurzzeit-Free-Recall-Test“ in entscheidendem Maß gefördert zu werden.
2. Auch unter der Free-Recall-Bedingung treten schlafbasierte Leistungssteigerungen auf. Im Gegensatz dazu sind diese nach einer Wachphase nicht zu beobachten. Hier kommt es lediglich zu einer Leistungsstabilisierung.
3. Da durch die Free-Recall-Bedingung in den Behaltenstests weder visuelle Instruktionen noch Feedback zur Verfügung stand, müssen die beobachteten Leistungsverbesserungen tatsächlich auf schlafbasierte Optimierungen an der Gedächtnisrepräsentation der neu erlernten Sequenz zurückzuführen sein. Ob es sich dabei um Optimierungen an der abstrakt-räumlichen Repräsentation oder einer motorischen Repräsentation handelt, muss dabei noch genauer untersucht werden.

4.4 Publikation 4: Task complexity modulates sleep-related offline learning in sequential motor skills (Blischke & Malangré, 2017)

4.4.1 Ziel

In Publikation 4 wurde untersucht, welche spezifischen Aufgabeneigenschaften dazu führen, dass die genannten Konsolidierungseffekte auftreten. In der jüngeren Vergangenheit sind Studien von verschiedenen Autoren (Genzel et al., 2012; Gudberg et al., 2015; Al-Sharman & Siengsukon, 2013, 2014) erschienen, die schlafbasierte Leistungssteigerungen bei komplexen, alltagsnahen Aufgaben untersucht haben. Allen genutzten Aufgaben ist eine gewisse Komplexität gemein.

Für den Begriff „Aufgabenkomplexität“ existiert keine allgemeingültige Definition (siehe Abschnitt 2.7.2 zu Komplexität). Allerdings kann von folgenden Faktoren angenommen werden, dass sie die Komplexität einer Aufgabe modulieren: Sequenzlänge, Distanz und Größe von Bewegungszielen, Regelmäßigkeit einer Sequenz, Anzahl von Endeffektoren, bilaterale Ausführungsbedingungen, Koordination zwischen verschiedenen Extremitäten bzw. Körperteilsegmenten, Redundanzkontrolle.

Überraschenderweise findet sich nur eine Veröffentlichung, die den Zusammenhang zwischen Aufgabenkomplexität und schlafbasierten Leistungsgewinnen systematisch untersucht (Kuriyama et al., 2004). Eine weitere Studie von Schmidt et al. (2010) untersuchte, ob schlafbasierte Leistungssteigerungen auch bei Ausführung der SRTT als Armbewegungssequenz mit einem oder beiden Händen auftreten. Dabei kam die originale 5-teilige SRTT zum Einsatz. Zudem kam in zwei separaten Untersuchungsgruppen die gleiche 5-teilige Sequenz als Armbewegungssequenz, die entweder mit einem oder zwei Endeffektoren ausgeführt wurde, zum Einsatz. Allerdings waren die Präzisionsanforderungen unter dieser Bedingung sehr gering. Im Gegensatz zur mit 4 Fingern ausgeführten SRTT finden sich unter den großmotorischen Bedingungen keine schlafbasierten Offline-Leistungssteigerungen. Bei gleicher Sequenzlänge (und damit gleich großem Informationsumfang der zu lernenden Sequenz) wurde die Anzahl der Endeffektoren variiert. Allerdings wurden gleichzeitig die Uni- bzw. Bilateralität der Aufgabenausführung sowie der komplette Endeffektor variiert (und damit die Notwendigkeit von großmotorischer Bewegungsproduktion), so dass keine

differenzierten Aussagen über den gesonderten Einfluss spezifischer Aspekte von Aufgabenkomplexität auf die Lernleistung möglich sind.

Der Zugang von Kuriyama et al. (2004) war etwas differenzierter: Sie untersuchten vier verschiedene Variationen der SRTT. Dabei wurden die Länge der Sequenz (und damit der zu encodierende Informationsumfang), die Uni- oder Bilateralität sowie die Anzahl der Endeffektoren variiert. Während allerdings bei einer Sequenzlänge von fünf Elementen jeweils insgesamt vier Finger (entweder vier einer Hand oder je zwei beider Hände) genutzt wurden, variierte die Zahl der genutzten Finger bei einer Sequenzlänge von neun Elementen: Hier wurden in der unimanualen Bedingung vier Finger genutzt. Beim Transfer auf die bimanuale Bedingung wurde allerdings gleichzeitig die Anzahl der genutzten Endeffektoren verändert. Zwar waren die gefundenen Effekte in dieser Bedingung am größten, allerdings konnte nicht differenziert werden, auf welchen die Aufgabenkomplexität beeinflussenden Faktor sie zurückzuführen sind: auf die bimanuale Ausführung oder auf die Erhöhung der Anzahl an Endeffektoren?

Insgesamt steht auf der Basis dieser beiden Veröffentlichungen zu vermuten, dass die Komplexität der verwendeten Aufgabe ein Einflussfaktor für das Auftreten bzw. für die Ausprägung von schlafbasierten Leistungssteigerungen ist. Die erhöhte Komplexität kommt dabei nicht allein durch die Zahl der verwendeten Endeffektoren, sondern durch die Kombination beider Hände mit jeweils fünf Fingern zustande. Bei jedem Tastendruck entstehen so 25 Kopplungsmöglichkeiten von Hand und Finger. Denn es muss nicht nur jeweils ein Finger aktiviert werden. Gleichzeitig muss auch die Aktivität der restlichen Finger gehemmt werden (siehe hierzu auch Yokoi et al., 2017).

4.4.2 Methode

In Publikation 4 wurde diese Hypothese unter Nutzung des Peg-Plugging-Paradigmas anhand von drei Experimenten untersucht. Im Gegensatz zu den oben genannten Studien (Schmidt et al., 2010; Kuriyama et al., 2004) wurde die Aufgabenkomplexität allerdings in Bezug auf nur zwei Komponenten systematisch variiert, um differenzierte Aussagen über jede der beiden Komplexitätskomponenten treffen zu können und Konfundierung wie in den oben genannten Studien zu vermeiden:

1. Die Länge der Sequenz (Anzahl von Sequenzelementen) und damit die „Größe“ bzw. den informationellen Umfang (also die Menge der zu speichernden Information) der abstrakt-räumlichen Repräsentation.
2. Die strukturelle Komplexität – „Regelhaftigkeit/Bekanntheit“ von Sequenzkonfigurationen – vs. fixe, aber nicht regelhaft strukturierte Sequenzkonfiguration. Für die „regelhafte“ Sequenz wurden Bewegungsendpunkte gewählt, die zusammen genommen bekannten geometrischen Formen, wie bspw. Drei- oder Vierecken, entsprachen.

Um die Störvarianz zu minimieren, wurden weitere die Sequenzausführungsleistung modulierende Faktoren – Fitts' ID (als Parameter für die Aufgabenschwierigkeit von Zielbewegungen mit Werten zwischen 4.95 und 5.35) oder die Anzahl der Endeffektoren (jeweils *ein* Arm) – über alle drei Experimente konstant gehalten. Somit konnte gewährleistet werden, dass einzig die Variation von Sequenzlänge und Regelhaftigkeit für unterschiedliche Ergebnisse infrage kamen.

Folgende Experimente wurden für Publikation 4 durchgeführt:

Experiment 1: Armbewegungssequenz bestehend aus 10 Elementen (entsprechend der in Publikation 3 verwendeten Sequenz)

Experiment 2: Reduktion auf fünf Sequenzelemente. Dabei handelte es sich bei den fünf Elementen um einen Ausschnitt aus der für Experiment 1 genutzten nicht regelhaften 10er-Sequenz, die ungefähr vergleichbare Transporttrajektorien beinhaltete.

Experiment 3: Armbewegungssequenz bestehend aus 10 Elementen, mit reduzierter struktureller Komplexität, operationalisiert über eine Sequenz, die regelhaft organisiert war, in Anlehnung an bekannte geometrische Muster (Dreiecke und Rechtecke).

Das Experimentalprotokoll beinhaltete – wie in den vorhergehenden Untersuchungen – jeweils zwei Untersuchungsgruppen („Morgen-Abend“ vs. „Abend-Morgen“) und jeweils ein 12-stündiges Behaltensintervall mit anschließendem Retest. Im Retest wurde stets dieselbe Stimulusinformation präsentiert wie in der Aneignungsphase. Die Zahl der Sequenzausführungen in der Aneignungsphase wurde wie in Publikation 1 auf 100 festgelegt. Diese 100 Trials waren wiederum in 10 Blöcke à 10 Trials aufgeteilt.

4.4.3 Ergebnisse

Die Ergebnisse von Publikation 4 zeigen, dass einzig in Experiment 1, also bei Verwendung einer nicht regelhaft strukturierten Armbewegungssequenz, schlafbasierte Leistungssteigerungen auftreten. Eine Reduktion der Sequenzlänge und damit der Informationsmenge bezüglich der abstrakt-räumlichen Bewegungsendpunkte in Experiment 2 führte lediglich zu einer Stabilisierung der Kriteriumsleistung unabhängig von Wachheit oder Schlaf. Das Erlernen der regelhaften Sequenz in Experiment 3 führte dagegen zu einer leichten Leistungssteigerung unabhängig von Wachheit oder Schlaf.

4.4.4 Weiterführende Überlegungen und Auswertungsschritte

Bei der Lektüre von Publikation 4 mag sich der Leser die Frage stellen, wie festgestellt wurde, ob eine „regelhafte“ Sequenz tatsächlich eine geringere Komplexität aufweist als eine Sequenz mit gleicher Anzahl an Elementen (und damit gleichem Informationsumfang) und einer fixen, aber keiner erkennbaren Produktionsregel folgenden und für die Probanden zu Beginn nicht vorhersagbaren Abfolge von Sequenzelementen. Die Antwort liegt in der Kombination aus drei Faktoren: Schwierigkeitsindex (Fitts' ID), Ausführungsdauer und Fehlermaß.

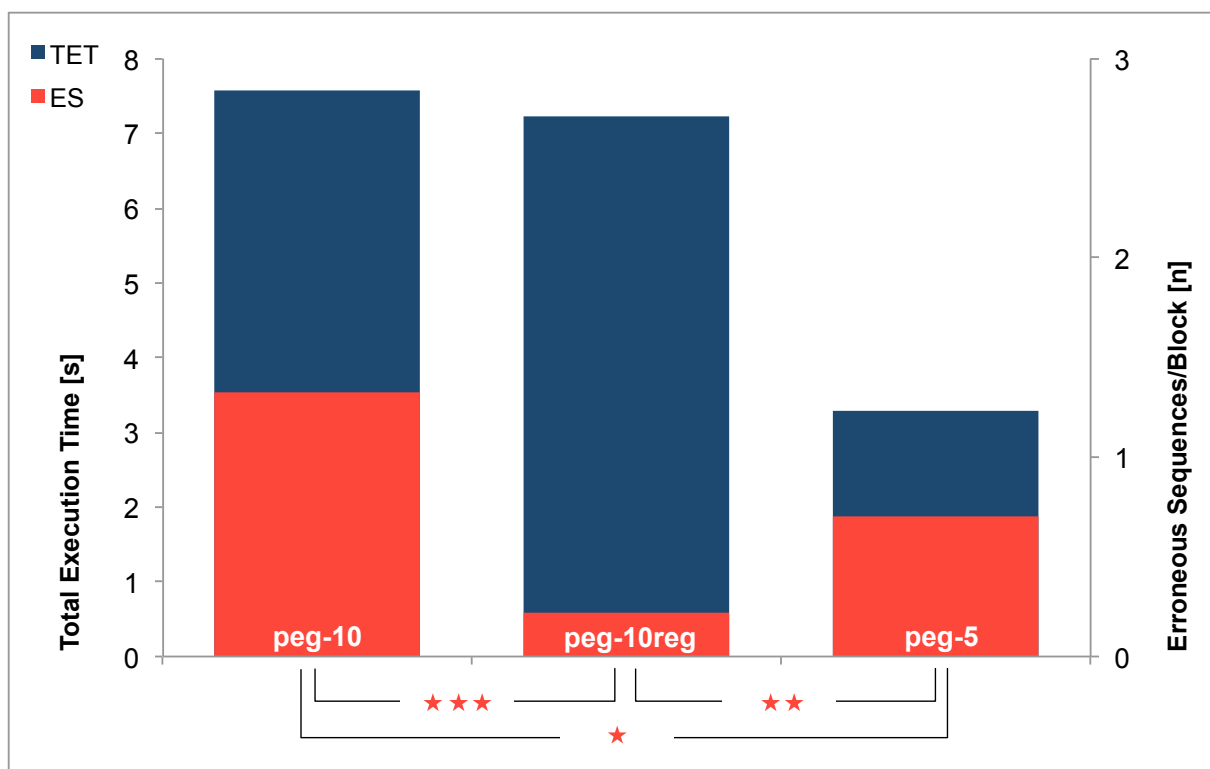


Abbildung 6: Dargestellt sind mit Total Execution Time (TET [s], dunkelblau) und fehlerhaften Sequenzen pro Block (ES [n], rot) Leistungs- und Fehlermaß für die drei Experimente aus Publikation 4 dargestellt. Die drei Sterne entsprechen $p < .001$ für den paarweisen Vergleich der Fehlerwerte zwischen der nicht regelhaften und der regelhaften Sequenz mit je zehn Elementen, die zwei Sterne: $p = .001$, der eine Stern $p = .011$. Abgebildet sind jeweils die Mittelwerte aus den Messzeitpunkten Post-Training und Retest.

Sollten die beiden Sequenzen mit je zehn Elementen (regelhaft vs. nicht regelhaft) die gleiche Komplexität aufweisen, wäre zu erwarten, dass bei ähnlichem Fitts' ID global gesehen eine ähnliche Ausführungsgeschwindigkeit bei fehlerfreier Ausführung und ähnliche Fehlerzahlen auftreten. Ein Blick auf Abbildung 6 zeigt, dass dieser Fall für die Ausführungsdauer pro Sequenz (TET) zutrifft. Allerdings zeigt sich ein signifikanter Unterschied zwischen den Fehlerwerten beider Experimente. Während bei der nicht regelhaften Sequenz im Mittel ca. 3,5 fehlerhafte Sequenzen pro Trial-Block auftreten, bewegt sich dieser Wert bei der regelhaften Sequenz nur knapp über null. Damit ist er sogar noch geringer als die durchschnittliche Fehlerrate der nicht regelhaften Sequenz mit nur fünf Elementen. Dieser Befund liefert einen deutlichen Hinweis darauf, dass die regelhafte Zehnersequenz eine niedrigere Komplexität aufweist als die nicht regelhafte.

Der Unterschied zwischen den Fehlermaßen der drei Experimente wurde anhand einer 2 (MZP: PT, RT) $\times 3$ (Experiment: 1, 2, 3) $\times 2$ (Retentionsbedingung: sleep, wake)-ANOVA überprüft. Es fanden sich kein signifikanter Haupteffekt MZP ($p = .073$) sowie keinerlei signifikante Zweifach- oder Dreifachinteraktionen. Ebenso zeigte sich kein signifikanter Haupteffekt Retentionsbedingung ($p = .247$) und kein Interaktionseffekt Retentionsbedingung \times Experiment ($p = .523$). Dagegen liegt ein signifikanter Haupteffekt „Experiment“ ($F_{[\text{Experiment}]}(1, 2) = 18,253$; $p < .001$; $\eta^2 = .353$) vor, der zudem eine klare Hierarchie der Fehlerraten erkennen lässt (Peg-10 (Mittelwert = 1,33) > Peg-5 (Mittelwert = 0,70) > Peg-10reg (Mittelwert = 0,22)).

Zur weiteren Analyse wurde ein Bonferroni-korrigierter Post-hoc-Test „Paarweise Vergleiche“ durchgeführt. Hier fanden sich signifikante Unterschiede zwischen allen drei Experimenten: Peg-10 vs. Peg-5 ($p = .001$), Peg-10 vs. Peg-10reg ($p < .001$) sowie Peg-5 vs. Peg-10reg ($p = .011$).

4.5 Zusammenfassung der vier Publikationen

Anhand der vier in den vorherigen Unterkapiteln beschriebenen Publikationen konnte nachgewiesen werden, dass a) schlafbasierte Leistungssteigerungen auch bei einer großmotorischen Armbewegungssequenz auftreten, b) schlafassoziierte Konsolidierungsprozesse für Optimierungen an der kognitiven Gedächtnisrepräsentation sorgen (unabhängig von schlafassoziierten Einflüssen auf den visuomotorischen Informationsumsatz im Falle von im Retest verfügbarer Stimulusinformation) und dass c) diese Optimierungen unter anderem auf Chunk-Concatenation – eine qualitative Optimierung der Gedächtnisrepräsentation – zurückzuführen sind. Des Weiteren wurde gezeigt, dass d) Aufgabenkomplexität, operationalisiert über Sequenzlänge einerseits und strukturelle Komplexität andererseits, ein modulierender Faktor für das Auftreten schlafbasierter Konsolidierungsprozesse, die zu Leistungszuwächsen führen, ist. Abbildung 7 zeigt die prozentualen Leistungsveränderungen in der Total Execution Time (TET) der einzelnen Experimente in den relevanten Behaltenszeiträumen, einmal nach dem jeweils durchwachten Behaltenszeitraum (grüne Balken) und einmal nach dem durchschlafenen Intervall (dunkelblaue Balken). Wiedergegeben sind jeweils die mittleren Change-Scores, die nach folgender Formel berechnet wurden:
$$\left[\left(\text{TET}_{\text{Post-Retention-Intervall}} - \text{TET}_{\text{Prä-Retention-Intervall}} \right) / \text{TET}_{\text{Prä-Retention-Intervall}} \right] * 100\% * (-1)$$
. Dieses Vorgehen stellte die Relativierung der Leistungsveränderung am Ausgangsniveau sicher und vermied so einen möglichen Bias der Leistungsveränderung. Im dem hier dargestellten Forschungsprogramm wurden also nicht nur global schlafbasierte Leistungsveränderungen bei einer großmotorischen Armbewegungssequenz festgestellt, sondern in einer differenzierten Herangehensweise auf verschiedenen Verhaltensebenen Einflussfaktoren für die gefundenen schlafbasierten Leistungssteigerungen bei dieser Aufgabe identifiziert. Dieses Vorgehen unterscheidet sich von dem anderer Forschungsgruppen, die in ihren Untersuchungen bisher meist lediglich einen Faktor, bspw. das Alter (siehe Gudberg et al., 2015) variiert haben.

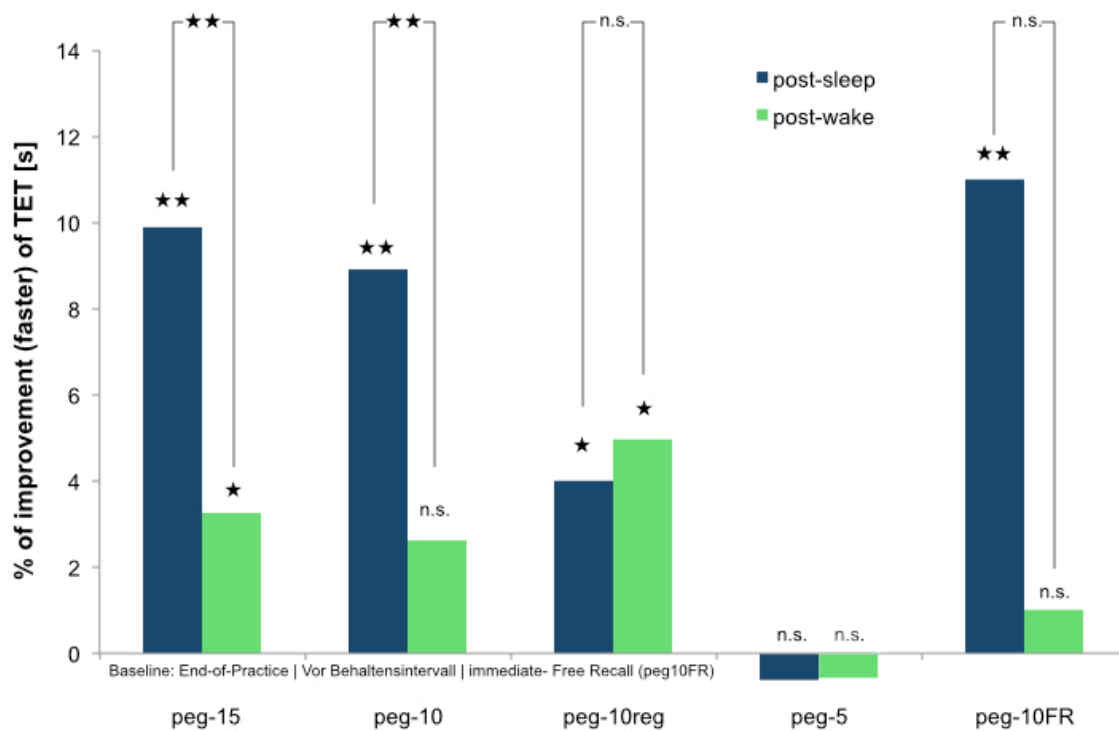


Abbildung 7a: Übersicht über die prozentualen Leistungsveränderungen (Leistungsgewinne) über den Schlafzeitraum (blaue Balken) und über den Wachzeitraum (grüne Balken) aller in den Publikationen 1 bis 4 beschriebenen Experimente.

Aufgrund der Nutzung aller zur Verfügung stehenden Daten beziehen sich die abgetragenen Werte für Peg-10, Peg-10reg und Peg-5 auf Leistungsveränderungen zwischen dem Ende der Aneignungsphase 3 und dem ersten Retentionstest nach 12 Stunden.

Bei Peg-15 sowie peg-10FR (FR = Free-Recall) sind die Daten aller Behaltensintervalle (1. und 2. 12-stündiges Behaltensintervall) enthalten. Abgetragen sind jeweils die über beide Gruppen gemittelten Veränderungen nach dem jeweiligen Schlaf- bzw. Wachintervall. Als Baseline für Peg-10FR dient der Early-Free-Recall-Test 15 Minuten nach Ende der Aneignungsphase. Die

Auflösung für jedes Behaltensintervall ist in Abbildung 7b zu sehen. Abkürzungen und Sterne siehe Fußnoten ⁴, ⁵ und ⁶.

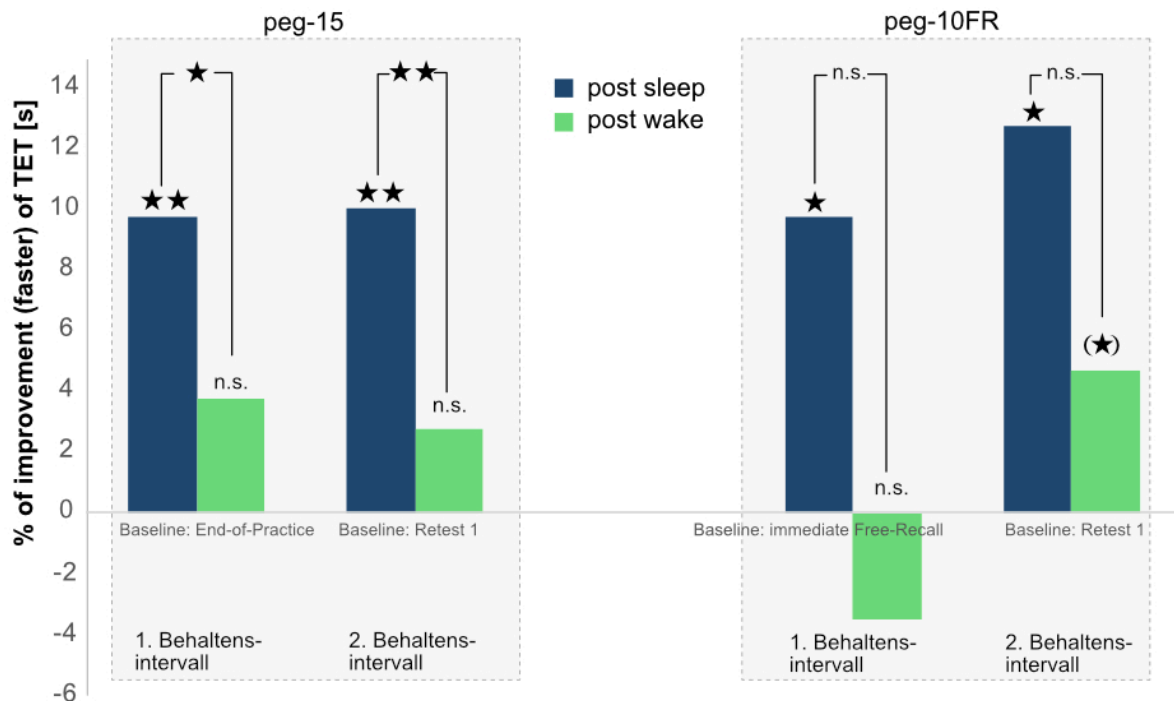


Abbildung 7b: Dargestellt sind die prozentualen Veränderungen aller Probanden nach dem 1. und 2. Behaltensintervall und nach Schlaf und Wachheit der Experimente Peg-15 sowie Peg-10FR. Baseline ist jeweils die Leistung vor dem jeweiligen Behaltensintervall.

Aus den Abbildungen 7a und 7b ergeben sich einige interessante Beobachtungen: Betrachtet man die verschiedenen Varianten der Peg-Plugging-Aufgabe, so scheint es

⁴ Sterne direkt über den Balken symbolisieren eine signifikante Veränderung gegenüber der Ausgangsleistung. Sterne zwischen zwei Balken symbolisieren einen signifikanten Unterschied zwischen den beiden Untersuchungsbedingungen (Wachheit vs. Schlaf); * $p < .05$, ** $p < .01$, (*) $p < .1$.

⁵ Abkürzungen: Peg-15: Peg-Plugging-Sequenz mit 15 in nicht regelhafter Abfolge angeordneten Elementen; Peg-10: Peg-Plugging-Sequenz mit 10 in nicht regelhafter Abfolge angeordneten Elementen; Peg-10reg: Peg-Plugging-Sequenz mit 10 regelhaft angeordneten Elementen; Peg-5: Peg-Plugging-Sequenz mit 5 in nicht regelhafter Abfolge angeordneten Elementen; Peg-10FR: Peg-Plugging-Sequenz mit 10 in nicht regelhafter Abfolge angeordneten Elementen und Test unter Free-Recall-Bedingungen.

⁶ Die entsprechenden TET-Change-Scores wurden allgemein nach folgender Formel berechnet: $[(TET_{\text{Post-Retention-Intervall}} - TET_{\text{Prä-Retention-Intervall}}) / TET_{\text{Prä-Retention-Intervall}}] * 100\% * (-1)$. Da hier eine Verringerung der Total Execution Time (TET) einen Leistungsgewinn repräsentiert, wurde die ermittelte, am Ausgangswert relativierte prozentuale TET-Differenz zwischen zwei Testzeitpunkten jeweils mit -1 multipliziert, um dem Betrachter einen unmittelbar intuitiven Eindruck von Leistungsverbesserungen (positiver Wert) und Leistungsverschlechterungen (negativer Wert) zu vermitteln.

keine lineare Beziehung zwischen der Aufgabenschwierigkeit und dem Ausmaß der schlafassoziierten Leistungssteigerungen hinsichtlich der Ausführungsgeschwindigkeit, operationalisiert über die TET, zu geben. Bei visueller Inspektion scheint es vielmehr so zu sein, dass die deutlichen Leistungsveränderungen nach Schlaf bei den „komplexeren“ Aufgabenvarianten Peg-15, Peg-10 sowie Peg-10FR auf ungefähr einem Level liegen. Bei den beiden weniger komplexen Aufgabenvarianten Peg-10reg und Peg-5 scheint es so, dass diejenige mit der größeren Sequenzlänge (Peg-10reg) über beide Intervalle geringe, aber gleich große, Peg-5 hingegen keinerlei Leistungsveränderungen mehr erfährt. Letzteres könnte allerdings auch auf einen Warm-up-Effekt in den Testblöcken zurückzuführen sein, der leichte Leistungssteigerungen nach den Behaltensintervallen verdeckt.

Des Weiteren ist sichtbar, dass bei allen Aufgabenvarianten mit zehn oder mehr Sequenzelementen auch nach Wachintervallen eine leichte Leistungsverbesserung im Bereich von 2,5 bis 5 % auftritt. Die Gründe hierfür sind bislang nicht bekannt. Infrage kommt, dass es sich dabei um eine implizite Komponente der Repräsentationsoptimierung handelt, deren Konsolidierungsprozesse eine zeitbasierte Leistungssteigerung zur Folge haben. Dies würde bedeuten, dass hier eher die motorische Ausführung der Aufgabe „Peg-Plugging“ – also der Armbewegung zu einem bestimmten Zielpunkt sowie des Einfädelns des Stiftes in das entsprechende Loch – optimiert wird als die Repräsentation der Sequenzabfolge. Das würde bedeuten, dass eine Verbesserung der Leistung auch bei einer randomisierten Sequenzabfolge auftreten müsste, was in zukünftigen Untersuchungen unter Umständen durch eine separate Kontrollgruppe zu prüfen ist.

Eine weitere Erklärung für die Leistungsgewinne nach Wachheit ist Weiterlernen im Retest in den Gruppen, bei denen die Retests unter Verfügbarkeit visueller Stimuli stattfanden. In den Free-Recall-Gruppen ist es denkbar, dass sich die Leistungssteigerungen durch das Weiterüben wiederum auf die motorische Aufgabenausführung und den Abruf der Sequenzrepräsentation beschränkte. Betrachtet man Abbildung 7b, entsteht der Eindruck, dass diese Funktion erst dann wirklich effektiv wird, wenn sie auf einer vergleichsweise sicheren Grundlage (bezüglich der Abfolge der auszuführenden Sequenz) stattfindet: nämlich dann, wenn das Wachintervall nach dem repräsentationsoptimierenden Schlafintervall liegt.

Eine weitere Frage zu den Ergebnissen von Peg-10FR ergibt sich aus Abbildung 7a: Warum ergibt sich trotz eines signifikanten Unterschieds der Kriteriumsleistung nach dem Schlafintervall und eines sich nicht signifikant von null unterscheidenden Change-Scores nach dem Wachintervall (und dem größten Unterschied der fünf Experimente bezüglich der deskriptiven Daten) kein signifikanter Unterschied zwischen Schlaf- und Wachintervall?

Zwei Erklärungsmöglichkeiten kommen aus Sicht des Autors hier infrage: Erstens waren aufgrund der Schwierigkeiten der Probanden im ersten 12-Stunden-Behaltensintervall die Sequenz korrekt zu reproduzieren, große Fluktuationen zu beobachten. In der MEM-Gruppe wurden die Leistungsver schlechterungen nach dem Wachintervall ($-3,5\%$ im 1. Behaltensintervall) im darauffolgenden Schlafintervall deutlich kompensiert ($+12,72\%$, siehe Abbildung 7b). Dies führte insgesamt zu sehr hohen Streuungswerten. Zweitens bestand diese Untersuchung aus zwei zeitlich versetzten Teiluntersuchungen mit unterschiedlichen Versuchsleitern bestand. Dies führte von Anfang an zu globalen und relativ stabilen Leistungsunterschieden zwischen beiden Gruppen. Auch dies trug zu den großen Streuungswerten bei.

5 Einige Überlegungen zur Auswertungsmethode

Ergebnisse der Publikationen 1 bis 4 erlauben erstmals einen detaillierten Einblick in das Thema schlafbasierter Leistungssteigerungen bei einer großmotorischen Aufgabe. Bisherige Forschungen widmeten sich lediglich den globalen Fragenstellungen nach schlafbasierten Offline-Leistungssteigerungen bei der Nutzung eines großmotorischen Paradigmas, gingen allerdings nicht weiter in die Tiefe (Genzel et al., 2012; Schmidt et al., 2010; Kempler & Richmond, 2012). Trotz der vergleichsweise detaillierten Auseinandersetzung mit dem Paradigma bleiben einige Fragen offen, die in zukünftigen Studien adressiert werden sollten. So stellt sich bspw. die Frage nach dem idealen methodischen Vorgehen für die Untersuchung von Konsolidierungsprozessen auf der Ebene verhaltensseitig beobachtbarer Leistungsveränderungen. Zudem ist die Generalisierung der Ergebnisse über verschiedene Populationsgruppen und Aufgabentypen sowie deren jeweilige motorische Kontrollanforderungen ungeklärt (siehe auch Gudberg et al., 2015).

Zunächst wird auf das Thema der methodischen Vorgehensweise eingegangen. In der jüngsten Vergangenheit hat sich aufgrund neuer Erkenntnisse eine Diskussion über das generelle methodische Vorgehen zur Evaluierung schlafbasierter Leistungssteigerungen entwickelt. Einige Publikationen haben hier Fragen aufgeworfen, die noch nicht abschließend beantwortet sind und daher Gegenstand zukünftiger Untersuchungen sein müssen. In den folgenden Kapiteln werden diese dort diskutierten Aspekte aufgeführt, und soweit möglich, erfolgt eine Einordnung der Publikationen 1 bis 4 bezüglich der jeweiligen Fragestellung.

5.1 Methodische Fragestellungen

Einige Publikationen (Rickard et al., 2008; Brawn et al., 2010; Nettersheim et al., 2016) präsentieren Ergebnisse, die grundsätzliche Fragen bezüglich der methodischen Herangehensweise an das Thema der schlafbasierten Konsolidierung motorischer Repräsentationen aufwerfen.

So wird die Leistung bei der Sequenzausführungsgeschwindigkeit (und damit die bislang fast in allen Studien verwendete abhängige Variable) möglicherweise durch verschiedene Phänomene mehr oder weniger stark verzerrt. Bislang werden folgende Phänomene diskutiert:

- Verzerrung durch Datenmittelung
- Leistungsver schlechterung im ersten Testblock: Warm-up-Decrements vs. partielles Vergessen
- Ermüdungseffekte während der Aneignungsphase
- Reaktive Hemmung innerhalb eines Blockes
- Early Performance-Boost: kurzzeitig nach Aneignungsende auftretende, transiente Leistungssteigerungen
- Time of Testing: tageszeitspezifischer Testzeitpunkt
- Testing-Effekt / Weiterlernen im Behaltenstest

5.1.1 Datenmittelung

In der jüngeren Vergangenheit haben detailliertere Untersuchungen des SRTT-Paradigmas dazu geführt, bisherige Ergebnisse aus einem anderen Blickwinkel zu

betrachten und diese teilweise in ein neues Licht zu rücken (siehe für einen Überblick Pan & Rickard, 2015). Daraus ergibt sich eine Reihe von Faktoren, die es in künftigen Untersuchungen zu kontrollieren gilt und die teilweise eine veränderte Herangehensweise erfordern. In den folgenden Unterabschnitten werden diese Faktoren vorgestellt. Zudem wird dargestellt, ob diese Faktoren auch bei der Peg-Plugging-Aufgabe berücksichtigt werden müssen bzw. wie bei der Auswertung der Daten damit umgegangen wurde.

5.1.1.1 Datenmittelung und Online-Learning können zu verzerrten Ergebnissen führen

Pan & Rickard (2015) heben in ihrem Überblicksbeitrag einen interessanten Aspekt für die Auswertung von Datenreihen hervor, der unter Umständen zu einer Verfälschung der Ergebnisse führen kann. Es handelt sich um die Berechnung von Mittelwerten aus mehreren Versuchsblöcken zum Vergleich zwischen zwei Testzeitpunkten. Was ursprünglich dazu gedacht war, die Power von Tests zu erhöhen und eine möglichst große Datenmenge in die statistische Auswertung einzubeziehen, kann laut Pan & Rickard (2015) zu einer Überschätzung von Unterschieden zwischen zwei Messzeitpunkten führen. Sie argumentieren in ihrem Überblicksbeitrag folgendermaßen: Die Datenmittelung über mehrere Blöcke kann aufgrund von weiterem Leistungszuwachs (typisch für den Leistungsverlauf in frühen Lernstadien) während der Aufgabenausführung („Online-Learning“) dazu führen, dass ein signifikanter Leistungsunterschied zwischen zwei Testzeitpunkten festgestellt wird. Diese auf Online-Learning basierenden Leistungsunterschiede können fälschlicherweise als Offline-Learning interpretiert werden, wenn zwischen den Messzeitpunkten ein Behaltensintervall liegt. Um einen direkten Vergleich zwischen gemittelten Daten an Aneignungsende und Retest zu vermeiden, schlagen sie vor, eine Potenzfunktion an den Verlauf der Lernkurve des Aneignungszeitraumes anzupassen. Aus dieser Funktion lässt sich die Leistungsentwicklung für die weitere Aufgabenausführung extrapolieren. Diese prognostizierten Daten können dann wiederum anstelle der Daten aus den letzten Aneignungsblöcken zur statistischen Auswertung herangezogen werden.

Für die statistische Auswertung aller fünf für die vorgestellten Publikationen durchgeführten Experimente wurde das vorgeschlagene Verfahren angewendet. Nur wenn sich ein signifikanter Unterschied zwischen prognostizierten Leistungsdaten und

Daten aus den jeweiligen Tests ergab, wurde dies als Indikator für Leistungssteigerungen basierend auf dem Treatment während des Behaltenszeitraumes gewertet.

Eine Ausnahme dazu bildet Publikation 3: Hier fanden alle Tests unter Free-Recall-Bedingungen statt. Die Sequenz musste somit vollständig memoriert werden und es waren keine visuellen Stimuli verfügbar. Aufgrund dieser Tatsache ist anzunehmen, dass die tatsächlichen Leistungsdaten in den Tests schlechter sein können als die aus den Aneignungsleistungen extrapolierten Daten. Denn diese Schätzungen simulieren den weiteren Verlauf des informationell gestützten Lernens der Aneignungsphase. Eine Interpretation des Vergleiches von Schätzdaten aus dem Aneignungsverlauf und tatsächlichen Daten aus den Tests ohne Verfügbarkeit visueller Stimuli scheint daher sachlich nicht vertretbar.

Im Folgenden werden weitere Aspekte vorgestellt, welche bei der Berechnung von Blockmittelwerten zu einer Über- oder Unterschätzung gefundener Effekte führen können und somit bei der Datenauswertung sowie der Planung von Experimentalprotokollen berücksichtigt werden sollten. Jeder Aspekt wird kurz vorgestellt und ggf. erfolgt eine Einordnung der Vorgehensweise für die Publikationen 1 bis 4. Wenn nötig, werden zusätzliche statistische Auswertungen zur Einordnung vorgenommen.

5.1.2 Warm-up-Decrements

Warm-up-Decrements sind ein gut dokumentiertes, bei einer Vielzahl von Bewegungsaufgaben auftretendes Phänomen: Nach einer Phase körperlicher Inaktivität kommt es bei der Wiederaufnahme einer Aktivität zu einer kurzen, vorübergehenden Leistungsver schlechterung gegenüber zuvor gezeigten Leistungen (siehe auch Schmidt & Wrisberg, 1971; Wrisberg & Anshel, 1993; Ajemian et al., 2010; Joseph et al., 2013; Nacson & Schmidt, 1971). Dabei mag der Warm-up-Effekt auf das Erreichen der optimalen „Betriebsbedingungen“ des menschlichen Bewegungsapparates oder auf andere Mechanismen zurückzuführen sein. Ajemian et al. (2010) schlugen bspw. vor, dass Warm-up-Effekte aufgrund der Notwendigkeit einer Rekalibrierung des sensomotorischen Systems zustande kommen. Allerdings existieren keine Untersuchungen, die diese Theorie belegen (Bizzi & Ajemian, 2015). Aus dem Vergleich von allgemeinen versus spezifischen Warm-ups kann jedoch geschlossen werden, dass nicht nur metabolische oder muskuläre Aufwärmeeffekte für den Warm-up-Effekt verantwortlich sind (Bizzi &

Ajemian, 2015). Die „activity-set“-Hypothese (Nacson & Schmidt, 1971) schlägt dagegen vor, dass Warm-up-Effekte aufgrund eines suboptimalen Aktivierungszustandes (zu hohes oder zu geringes Aktivierungsniveau) auftreten. Erst mit dem Erreichen des optimalen Aktivierungsniveaus kann die Leistung erreicht werden, die dem tatsächlichen Fertigniveau entspricht. Rickard et al. (2008) mutmaßen, dass eine Leistungsverlechterung im ersten Testblock auf partielles Vergessen zurückzuführen sein könnte.

Da Warm-up-Decrements allerdings sowohl bei hohen Expertise-Levels als auch schon nach kurzen Zeiträumen der Inaktivität auftreten (Ajemian et al., 2010), kann davon ausgegangen werden, dass es sich bei transienten Leistungsverlechterungen zu Beginn eines Retests eher um Warm-up-Effekte denn um partielles Vergessen handelt. Diese These wird von Veröffentlichungen gestützt, die sich mit der verhaltensseitig beobachtbaren Leistung im zeitlichen Verlauf des motorischen Lernens beschäftigen (Joseph et al., 2013). Hier zeigt sich, dass der langfristige Leistungsverlauf bei der Aneignung einer motorischen Aufgabe (über mehrere Tage) am besten durch eine Exponentialfunktion oder Potenzfunktion beschrieben werden kann. Welche der beiden Funktionstypen den Verlauf am besten charakterisiert, hängt dabei unter anderem von der Art der Lernaufgabe und den daraus resultierenden Daten ab (Joseph et al., 2013). Diese Autoren fanden auch, dass der Exponent der Warm-up-Leistungskurve eine Größenordnung größer ist als der der langfristigen Leistungskurve. Die Autoren schlussfolgern daraus, dass Warm-up und langfristiger Leistungsverlauf unterschiedlichen Prozessen zuzuordnen sind. Leistungseinschränkungen zu Beginn einer neuen Übungssitzung sind laut ihnen dem Verlust spezifischer aufgabenrelevanter Systemanpassung oder noch mangelhafter allgemeiner Adjustierung von Aufmerksamkeit und physiologischen Systemzuständen zuzuschreiben (Joseph et al., 2013). Im langfristigen Leistungsverlauf finden sich also keine Anzeichen dafür, dass „partielles“ Vergessen auftreten sollte, denn dieses sollte durch deutlich länger dauernde Leistungsverluste gekennzeichnet sein.

Warm-up-Decrements können bei verschiedenen experimentellen Designs durch Datenmittelung dazu führen, dass die tatsächliche Leistung zu den Testzeitpunkten unterschätzt wird. Daher kommt dem Design der Retests und dem Umgang mit dem Warm-up-Effekt in der Konsolidierungsforschung ein besonderer Stellenwert zu. Bislang existieren – bezogen auf die Erforschung von Konsolidierungseffekten – nur wenige

Stellen (bspw. Reis et al., 2015), welche das Phänomen erwähnen und diskutieren. Eine systematische Auseinandersetzung über den Umgang mit dem Phänomen „Warm-up“ findet jedoch in der Schlaf- bzw. Konsolidierungsforschung nicht statt.

5.1.2.1 Umgang mit Warm-up-Decrements in der vorliegenden Arbeit

In allen Experimenten der vorliegenden Arbeit tauchte das Phänomen der transienten Leistungsver schlechterung zu Beginn jedes Retests auf: Bei den ersten Sequenzausführungen des jeweils ersten Testblockes zeigten die Probanden durchweg schlechtere Leistungen als in den darauffolgenden Trials und in den darauffolgenden Blöcken, insbesondere aber auch im Vergleich zu den im oben angeführten Verfahren der aus den Aneignungsleistungen extrapolierten Leistungsdaten für die Testblöcke. In diversen Pilotstudien war das Phänomen ebenso zu beobachten. Mit dem Auftreten nach längeren Zeiträumen der Inaktivität (nach den Behaltensintervallen), nicht aber nach kurzen Zeiträumen (30-Sekunden-Intervalle zwischen den Blöcken) deckt sich diese transiente Leistungsver schlechterung mit den in der Literatur beschriebenen Warm-up-Decrements (Schmidt & Wrisberg, 1971; Wrisberg & Anshel, 1993; Ajemian et al., 2010; Joseph et al., 2013; Nacson & Schmidt, 1971).

Zwei mögliche Erklärungen für den Warm-up-Effekt ergeben sich aus der Literatur: einerseits die Notwendigkeit der Rekalibrierung des sensomotorischen Systems (Ajemian et al., 2010), andererseits die „activity-set“-Hypothese (Nacson & Schmidt, 1971).

Eine dritte plausible Erklärung ist es, dass die Probanden sich über ihre Kapazität zur sicheren Reproduktion der Sequenz, vor allem der räumlichen Bewegungsziele, zu Beginn der Tests nach einem 12-stündigen Behaltensintervall noch unsicher waren. Um die Sequenz, wie gefordert, fehlerfrei auszuführen, wurden zu Beginn langsamere Trials in Kauf genommen.

Zum jetzigen Zeitpunkt kann anhand der verfügbaren Daten allerdings keine verlässliche Aussage darüber getroffen werden, welche dieser drei Mechanismen – oder welche Kombination der Mechanismen – tatsächlich zutrifft. Alle drei würden das Phänomen der Warm-up-Effekte plausibel erklären. Um zukünftig Klarheit darüber zu erlangen, sind weitere Untersuchungen vonnöten.

Auch in weiteren Publikationen (bspw. Rångtjell et al., 2017; Kuriyama et al., 2004; Rickard et al., 2008; Fogel et al., 2017) drängt sich bei visueller Inspektion der Daten der

Verdacht auf, dass Warm-up-Decrements auftraten. Allerdings werden das Auftreten des Phänomens sowie der Umgang damit von den Autoren der Studien nicht weiter erläutert. Es finden sich jedoch in vereinzelt Untersuchungen Hinweise darauf, dass Warm-up eine Rolle spielt. Fischer et al. (2005) schließen bspw. in einer fMRI-Studie die ersten 6 Sekunden eines Übungsblockes von der Analyse aus, um nur den Zeitraum der höchsten Performance in die Auswertung einfließen zu lassen.

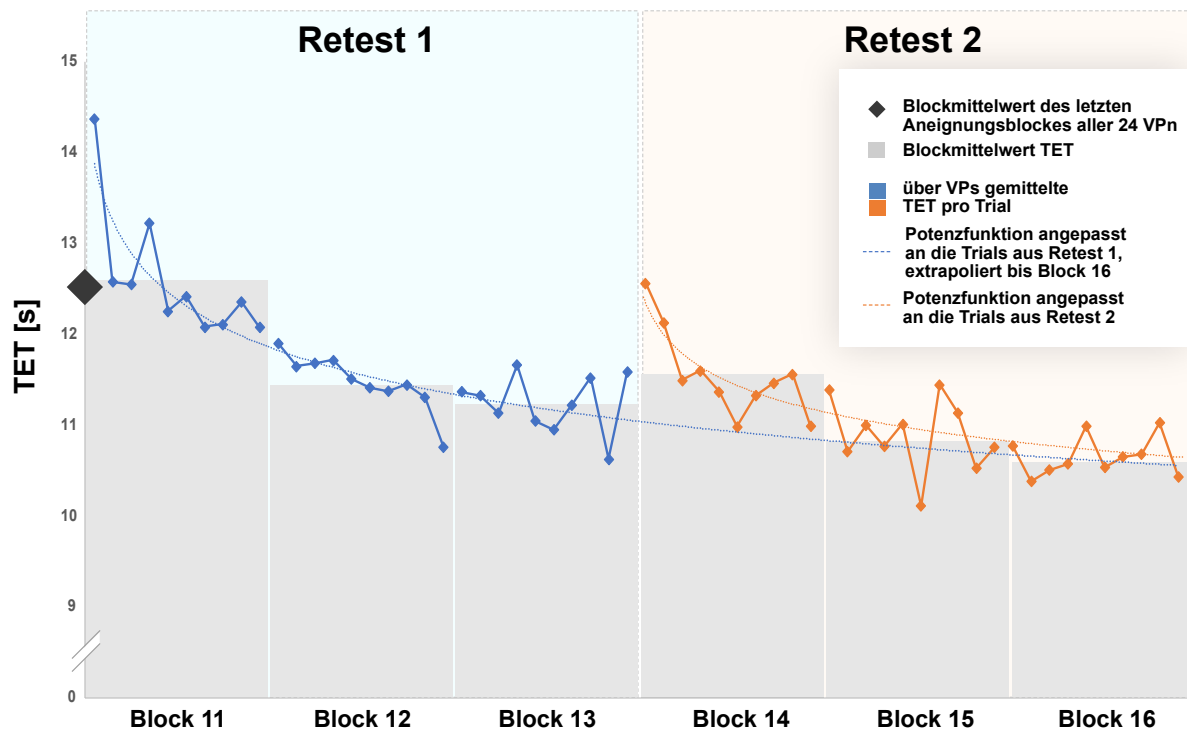


Abbildung 8: Beispielhafte Darstellung des Warm-up-Decrements anhand der über alle VPs gemittelten Einzelsequenzen der sechs Blöcke aus den Retests der Peg-Plugging-Aufgabe mit 15 Sequenzelementen (Publikation 1).

Für die Auswertung der Daten der vorgelegten Publikationen wurde folgende konservative Herangehensweise gewählt:

Alle Experimente enthielten jeweils drei Testblöcke, um den Einfluss von Warm-up-Effekten durch die Mittelung der Daten über drei statt zwei Blöcke gering zu halten. Alle drei Testblöcke wurden in die statistische Auswertung einbezogen. Somit wurde in Kauf genommen, dass die Daten der Tests leicht zuungunsten der tatsächlichen Leistungssteigerung verzerrt wurden.

Die Warm-up-Effekte traten unter allen Untersuchungsbedingungen auf (nach Wachheit ebenso wie nach Schlaf), so dass Treatment-bedingte Leistungsunterschiede zwischen diesen beiden Bedingungen jeweils unabhängig von Warm-up-Effekten auftreten sollten. Abbildung 8 zeigt beispielhaft den Verlauf der Leistungsentwicklung in beiden Testblöcken.

Anhand der TET-Daten von Publikation 1 (Peg-15) wurde zudem exemplarisch überprüft, ob der beobachtete Warm-up-Effekt in jedem Test auftritt und, wenn ja, welche Blöcke oder Sequenzen ihn am ehesten reflektieren. Hierzu wurde eine 10 (Sequenz: 1–10) \times 3 (Block: 1, 2, 3) \times 2 (Test: Retest 1, Retest 2)-ANOVA berechnet. Sie beinhaltete die Daten aller verfügbaren fehlerfreien Sequenzen aller Versuchspersonen in beiden Retests, gemittelt über beide Experimentalgruppen (vgl. Abbildung 8). Aufgrund diverser unsystematisch auftretender fehlender Werte (fehlerhafte Sequenzen oder technisch bedingte Ausfälle während der Datenaufnahme) auf individueller Ebene konnten die Daten allerdings nicht als abhängige Daten organisiert werden. Im Ergebnis zeigten sich keinerlei signifikante Interaktionseffekte. Dafür fanden sich signifikante Haupteffekte: „Sequenz“ ($F(9, 993) = 2,495; p = .008; \eta_p^2 = .022$), „Block“ ($F(2, 993) = 33,462; p < .001; \eta_p^2 = .063$) sowie „Test“ ($F(1, 993) = 34,715; p < .001; \eta_p^2 = .034$). Der Tukey-Post-hoc-Test zeigte signifikante Leistungsveränderungen zwischen dem ersten und zweiten Testblock ($p < .001$) sowie zwischen dem ersten und dritten Testblock ($p < .001$). Zwischen dem zweiten und dritten Testblock ergab sich keine signifikante Leistungsveränderung ($p = .178$). Die Ausführungsdauer im ersten Testblock war also jeweils signifikant langsamer als im zweiten und dritten Testblock. Für den Vergleich der Sequenzen wurden mithilfe des Tukey-Post-hoc-Tests zwei homogene Untergruppen gebildet. Hier zeigte sich, dass alle außer drei Sequenzen in beiden Untergruppen vertreten waren – die Sequenzen 6 und 10 mit sehr kurzen Zeitdauern und die 1. Sequenz mit einer sehr langen Zeitdauer. Das gesamte Ergebnis wurde folgendermaßen interpretiert: Für den Warm-up-Effekt scheint maßgeblich der 1. Ausführungsblock eines Retests verantwortlich zu sein. Gleichzeitig scheint die jeweils erste Sequenz sich von allen anderen Sequenzen abzuheben.

Auf der Basis dieser Erkenntnisse wurde eine zweite ANOVA berechnet. Hier wurde im Vergleich zur oben beschriebenen Berechnung lediglich die erste Sequenz des jeweils ersten Testblockes von der Berechnung ausgeschlossen. Die Ergebnisse zeigen, dass es unter diesen Bedingungen keinen signifikanten Haupteffekt „Sequenz“ mehr gab

($p = .336$). Die Haupteffekte „Block“ ($F(2, 965) = 24,184$; $p < .001$; $\eta_p^2 = .048$) und „Test“ ($F(1, 965) = 31,056$; $p < .001$; $\eta_p^2 = .031$) blieben allerdings signifikant. Der Tukey-Post-hoc-Test ergab nunmehr nur noch eine einzige homogene Untergruppe. Dieses Ergebnis zeigt, dass zu jedem Testzeitpunkt ein Warm-up-Effekt auftrat, den auf der Ebene einzelner Sequenzen vor allem die erste Sequenzausführung eines jeden Testzeitpunktes reflektierte. Auf der Ebene der Ausführungsblocks zeigte sich der Warm-up-Effekt vor allem dadurch, dass die Leistung im ersten Block signifikant schlechter war als in den beiden letzten Testblöcken, die sich statistisch nicht voneinander unterschieden.

Um zu überprüfen, ob der gefundene Warm-up-Effekt einen Einfluss auf die Ergebnisse der inferenzstatistischen Auswertung hat, wurde in allen drei Experimenten von Publikation 4 jeweils die TET-Leistungsveränderungen nur unter Berücksichtigung der beiden letzten Aneignungsblöcke (Blöcke 9 und 10) und der jeweils beiden letzten (Warm-up-freien) Testblöcke (Blöcke 12 und 13) inferenzstatistisch geprüft. Hinsichtlich der unterschiedlichen Wirkung von Wachheit vs. Schlaf ergaben sich hier keine abweichenden Ergebnisse gegenüber den Berechnungen unter Einbeziehung aller drei Testblöcke (siehe hierzu Blischke & Malangré, 2017; Supplemental Material).

Auch mit den Daten von Publikation 1 (Peg-15) wurde solch eine Überprüfung vorgenommen, hier auch unter Einbeziehung des zweiten Retests (i. e. Testblöcke 15 und 16). Hier zeigten sich ebenso wenig Unterschiede bezüglich der Ergebnisse, weder für die MEM-Gruppe noch für die EME-Gruppe: Für die MEM-Gruppe war der Haupteffekt „Testzeitpunkt“ ($F(2, 22) = 18,704$; $p < .001$; $\eta_p^2 = .630$) nach wie vor signifikant. Bei den Innersubjekt-Kontrasten ergab sich hier ein nur für das durchgeschlafene Behaltensintervall signifikanter Unterschied ($F(1, 11) = 22,021$; $p = .001$; $\eta_p^2 = .667$), während ein Tendenzeffekt für den Wachzeitraum zu erkennen war (PT – T1: $F(1, 11) = 4,030$; $p = .070$; $\eta_p^2 = .268$). Für die Gruppe EME ergab sich ebenso ein signifikanter Haupteffekt „Testzeitpunkt“ ($F(2, 22) = 28,676$; $p < .001$; $\eta_p^2 = .723$) und bei den Innersubjekt-Kontrasten zeigte sich hier nur für den durchgeschlafenen Zeitraum ein signifikanter Leistungsunterschied ($F(1, 11) = 65,856$; $p < .001$; $\eta_p^2 = .857$), nicht aber nach dem durchwachten Behaltensintervall ($F(1, 11) = 1,009$; $p = .337$; $\eta_p^2 = .084$).

Ob Warm-up-Effekte nun aufgrund der Aktivierung physischer Parameter, der „activity-set“-Hypothese (Nacson & Schmidt, 1971), der Rekalibrierungshypothese (Ajemian, 2010) oder aufgrund der oben vorgeschlagenen Möglichkeit der Rückversicherungsstrategie für fehlerfreie Aufgabenausführung auftreten: Für die

Zukunft wären eine systematische Untersuchung der Effekte bei verschiedenen Paradigmen und darauf aufbauend eine einheitliche Vorgehensweise für Experimente erstrebenswert, ggf. in Anlehnung an Ajemian et al. (2013).

5.1.3 Reaktive Hemmung

Rickard et al. (2008) berichten, dass während massierter Übung in der Aneignungsphase der SRTT mehrere Phänomene auftreten, welche die beobachtete Leistung verzerren. Sie fanden zu Beginn der Trainingsphase in jedem Block (30 Sekunden kontinuierliche Sequenzwiederholungen ohne Pause) einen typischen Leistungsverlauf: Die jeweils erste Sequenzausführung eines Blockes war die langsamste, darauf folgten durchgehend schnellere Sequenzausführungen. Ab dem sechsten Block fand sich allerdings ein abweichendes Muster: Die Sequenzen 5 bis 12 innerhalb eines Blockes wurden zunehmend langsamer. Ursache dafür sind laut Rickard et al. (2008) reaktive Hemmungseffekte (eine Leistungsver schlechterung zum Ende jedes 30-sekündigen Ausführungsblockes bei kontinuierlich ausgeführten Fingerbewegungssequenzen) sowie das Auftreten von Ermüdung bei massierter Übung der SRTT. Das führt bei Mittelung aller Daten pro Block zu einer Unterschätzung der tatsächlichen Leistungsfähigkeit am Ende der Aneignungsphase. Dies wiederum kann dazu führen, dass zwischen Aneignungsende und Retest ein Leistungsunterschied entsteht, der mehr auf die Datenmittelung der von reaktiver Hemmung geprägten letzten Aneignungsblöcke denn auf Veränderungen der Sequenzrepräsentation zurückzuführen ist.

Um dem entgegenzuwirken, wurden von Rickard et al. (2008) zur Berechnung der Leistung am Ende der Aneignungsphase nur die Sequenzen 2 bis 4 der letzten beiden Trainingsblöcke herangezogen. Die Ergebnisse veränderten sich dadurch dramatisch: Es traten nun keine schlafbasierten Leistungssteigerungen und keine Stabilisierung nach Wachheit mehr auf. Stattdessen stabilisierte sich die Leistung nach Schlaf und verringerte sich nach Wachheit.

Allerdings ergeben sich aufgrund der Darstellung von Rickard et al. (2008) durchaus einige Fragen, die von der Publikation nicht beantwortet werden und Gegenstand zukünftiger Untersuchungen sein müssen: Laut Aussage der Autoren führten die Probanden im Schnitt 22,3 bis 24,1 Sequenzen pro Block aus. Allerdings kamen für die Auswertungen nur zwölf davon zur Sprache. Wie sich die Leistung für die restlichen Sequenzausführungen darstellte, wird leider nicht dargestellt. Zudem wird nicht erklärt,

ob die Sequenzen 2 bis 4 nur für die Aneignungsblöcke gewählt wurden oder auch für die Testblöcke. Ob im ersten Block des Tests Warm-up-Effekte auftraten, kommt ebenfalls nicht zur Sprache.

In einer weiteren Variante versuchten Rickard et al. (2008), durch die Einführung von „spaced practice“ statt „massed practice“ reaktiven Hemmungs- und Ermüdungseffekten während der Aneignungsphase entgegenzuwirken.

Reaktive Hemmung im Peg-Plugging-Paradigma

Aufgrund der Ergebnisse der oben genannten Studien ist auch eine Auswertung der Daten des Peg-Plugging-Paradigmas bezüglich reaktiver Hemmungseffekte angebracht. Allerdings muss in Betracht gezogen werden, dass es sich beim Peg-Plugging im Gegensatz zur SRTT um eine eher diskrete Sequenzproduktionsaufgabe handelt (jeweils kurze Pausen zwischen dem Abschluss einer Sequenzausführung und dem Start der nächsten Ausführung auf der einen Seite sowie der Start der Bewegungsfolge erst nach dem Bereit-Signal der Probanden auf der anderen).

In Abbildung 9 ist die TET-Leistung der 1. bis 10. Sequenzausführung gemittelt über die Blöcke 8, 9 und 10 aller Probanden der Peg-15-Aufgabe dargestellt. Einige Anmerkungen dazu:

1. Zwar verlangsamte sich im Mittel die Ausführungsgeschwindigkeit von der 8. bis zur 10. Sequenzausführung geringfügig. Allerdings müssen diese Veränderungen in Anbetracht der geringen Mittelwertsunterschiede bei gleichzeitig erheblichen Streuungen als statistisch irrelevant betrachtet werden.
2. Sequenzausführung Nummer 8 war eine der schnellsten innerhalb des Blockes, somit kann von reaktiver Hemmung nicht die Rede sein.
3. Die Beschränkung der Datenauswertung auf die jeweils schnellsten Sequenzausführungen ist nicht empfehlenswert, da so nur ein kleiner Teil der verfügbaren Daten in die statistische Auswertung einflösse.

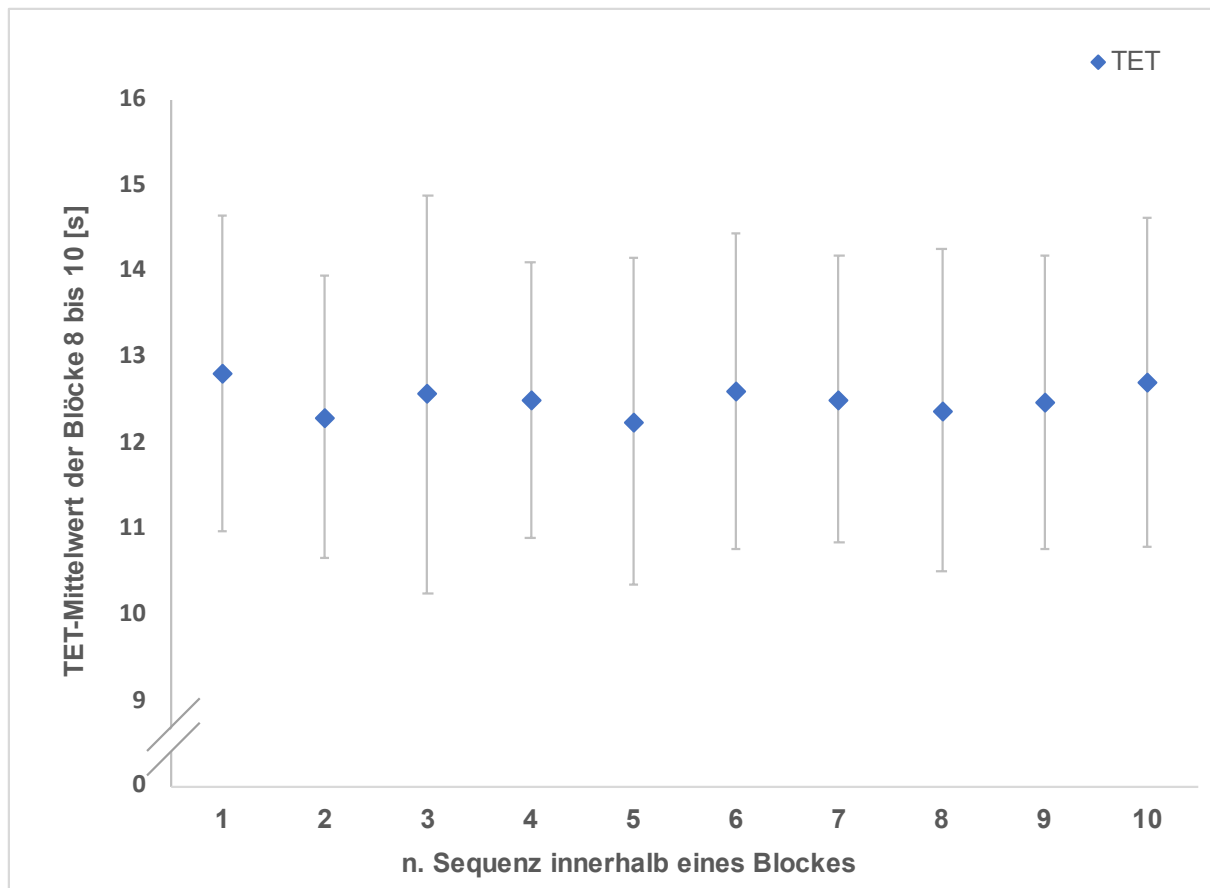


Abbildung 9: Wiedergegeben sind die Mittelwerte aller Probanden der Aneignungsblöcke 8 bis 10 (Rauten) und die Standardabweichungen (Fehlerindikatoren) der Total Execution Time der einzelnen Sequenzausführungen (Publikation 1 – Peg-15). Bei visueller Inspektion wird deutlich, dass alle Sequenzausführungen auf praktisch gleichem Niveau zu liegen scheinen.

Für die drei Experimente von Publikation 4 wurde zudem exemplarisch ermittelt, ob sich die jeweils ersten drei fehlerfreien Sequenzausführungen der Aneignungsblöcke 8, 9 und 10 von den drei jeweils letzten unterschieden. Falls reaktive Hemmung bei der Peg-Plugging-Aufgabe auftrat, sollte die Ausführung der drei letzten Sequenzen pro Block signifikant langsamer gewesen sein als die Ausführung der ersten drei Sequenzen pro Block.

Hierzu wurde eine 3 (Experiment: 1, 2, 3) \times 3 (Block: 8, 9, 10) \times 2 (Gruppe: EM, ME) \times 2 (Sequenzausführungen: erste drei, letzte drei)-ANOVA gerechnet. So wurde sichergestellt, dass die Daten aller Gruppen aus allen drei Experimenten für die relevanten Aneignungsblöcke genutzt werden. Die Sequenzausführungen „erste drei“ beziehen sich jeweils auf die ersten drei fehlerfreien Sequenzausführungen in jedem Block (Sequenzen 1, 2 und 3), die Sequenzausführungen „letzte drei“ auf die letzten drei

fehlerfreien in jedem Block (Sequenzen 8, 9 und 10). Die Ergebnisse zeigen einen signifikanten Haupteffekt „Block“ ($F_{\text{Block}}(1, 76, 117,56) = 12,134; p < .001; \eta^2 = .153$), die TET nahm im Mittel über die Blöcke 8, 9 und 10 ab. Ebenso fand sich ein signifikanter Haupteffekt „Gruppe“ ($F_{\text{Gruppe}}(1, 67) = 4,562; p = .036; \eta^2 = .064$). Dieser Effekt trat aufgrund der durchgehend kürzeren TET der ME-Gruppen auf. Der Haupteffekt „Experiment“ ($F_{\text{Experiment}}(2, 67) = 178,922; p < .001; \eta^2 = .842$) war zu erwarten, da sich die TET der Experimente zwangsläufig voneinander unterschied. Eine signifikante Interaktion „Gruppe \times Experiment“ fand sich nicht ($p = .161$).

Es zeigte sich weder ein signifikanter Haupteffekt „Sequenzausführungen“ ($p = .521$) noch ein signifikanter Interaktionseffekt dieses Faktors mit einem oder mehreren der anderen Faktoren ($p \geq .114$). Das bedeutet, dass sich die Ausführungsdauern für die ersten drei Sequenzausführungen der letzten Aneignungsblöcke nicht von denen der letzten drei Sequenzausführungen unterschieden (siehe auch Abbildung 10).

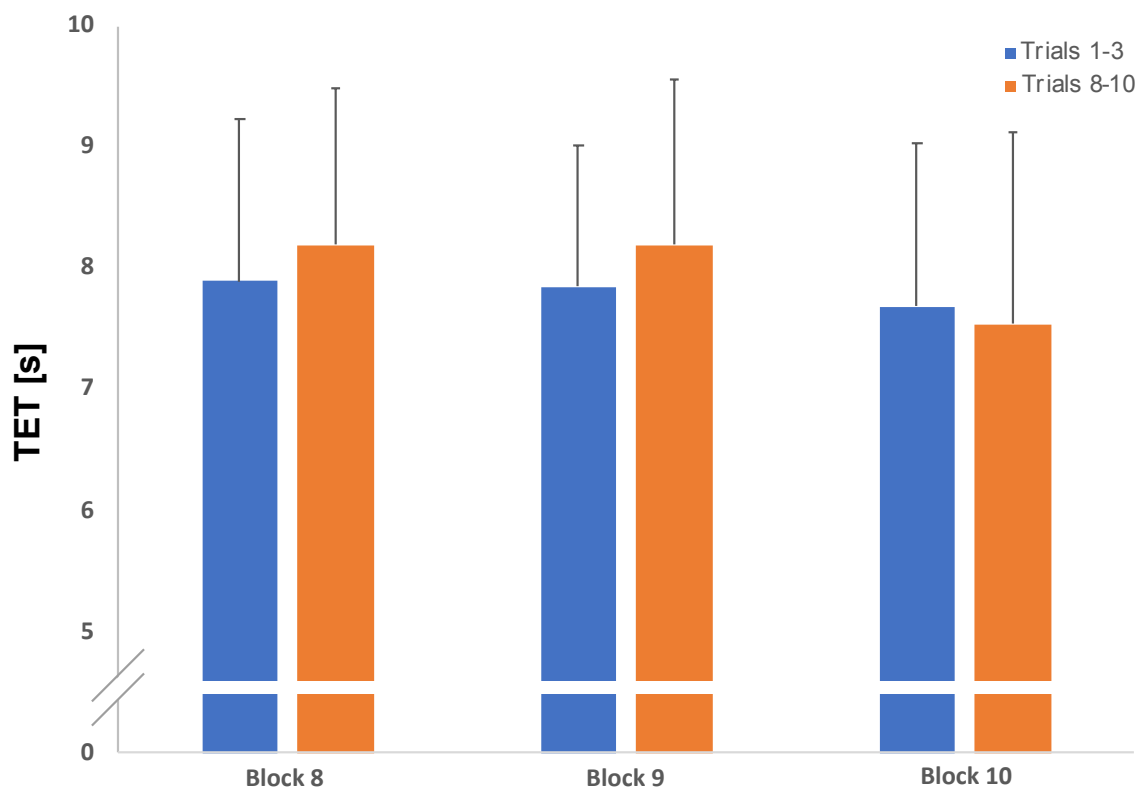


Abbildung 10: Mittelwerte (Balken) und Standardabweichungen (Fehlerindikatoren) der Total Execution Time (TET) aller Versuchspersonen der Experimente 1 bis 3 von Publikation 4. Die blauen Balken stellen die ersten drei Sequenzausführungen (Trials) des jeweiligen Blockes dar, die orangen Balken die letzten drei des jeweiligen Blockes.

Im Gegensatz zur SRTT und den Ergebnissen von Rickard et al. (2008) kann davon ausgegangen werden, dass reaktive Hemmungsprozesse beim Peg-Plugging-Paradigma keine Rolle spielen.

5.1.4 Early Performance Boost und reaktive Hemmung?

Hotermans et al. (2006), Brawn et al. (2010) sowie Nettersheim et al. (2015) zeigten, dass eine kurze Pause von fünf bis 30 Minuten ausreichend ist, um die Effekte von reaktiver Hemmung zu reduzieren: Bei einem Test innerhalb dieses Zeitraumes im Anschluss an die Aneignungsphase tritt eine gegenüber dem Aneignungsende signifikante Leistungssteigerung auf. Dieser Effekt wird von verschiedenen Autoren berichtet und als „Early Boost“ bezeichnet. Ob er allein auf die Auflösung von reaktiver Hemmung zurückzuführen ist, ist jedoch noch nicht geklärt. Nettersheim et al. (2015) sowie Hotermans et al. (2006) berichten ebenso, dass der beobachtete Early Boost eine transiente Erscheinung ist und vier Stunden nach Ende der Aneignung nicht mehr auftritt. Welche Mechanismen genau hinter dem Early Boost stehen, konnten bisherige Untersuchungen nicht beantworten. Pan & Rickard (2015) mutmaßen, dass der Early Boost durch den Abbau von Ermüdung oder reaktiver Hemmung zustande kommt und somit die tatsächliche Leistungsfähigkeit am Ende der Aneignungsphase darstellt. Allerdings könnten auch andere Erklärungen für das Auftreten des Early Boosts herhalten, bspw. eine frühe aktive Konsolidierungsphase. Interessant wäre in diesem Zusammenhang auch, ob eine qualitative Veränderung der Aufgabenausführung auftritt, z.B. Chunk-Concatenation.

Ob der Early Boost auch bei dem für die vier Publikationen genutzten Peg-Plugging-Paradigma auftritt, ist bislang noch nicht bekannt und sollte Gegenstand künftiger Untersuchungen sein. Einen Hinweis darauf könnten lediglich die Ergebnisse des ersten Behaltenstests, 15 Minuten nach Aneignungsende, aus Publikation 3 („Free Recall“) geben. Hier zeigte sich im Gegensatz zum Aneignungsende keinerlei Leistungsveränderung. Da diese Leistungsdaten allerdings unter anderen Instruktionsbedingungen (keine visuellen Stimuli) als die Leistungsdaten der Aneignungsphase erhoben wurden, kann hier keine eindeutige Aussage gemacht werden.

5.1.5 Ermüdung

Pereira et al. (2015) führten aufbauend auf den Ergebnissen und der Argumentation von Rickard et al. (2008) eine Studie durch, um die Art der Datenmittelung bei der bekannten Fingertapping-Task zu verändern und weitere Erkenntnisse über den tatsächlichen Betrag schlafbasierter Leistungssteigerungen zu gewinnen. Um unkontrollierte Varianz in den Daten zu verringern und um zu verhindern, dass die letzten Blöcke der Aneignungsphase durch Ermüdung oder Langeweile verzerrt werden, wurden statt der letzten drei Aneignungsblöcke jeweils die drei besten Aneignungsblöcke pro Versuchsperson für den Vergleich mit den Retestdaten herangezogen. So konnte gezeigt werden, dass sich die Ergebnisse beider Berechnungen signifikant voneinander unterscheiden: Bei der herkömmlichen Variante erscheinen die Leistungsgewinne durch Schlaf größer als bei der neu eingeführten Variante. Obwohl die Leistungsgewinne abnehmen, zeigt sich auch bei der neuen, alternativen Auswertungsmethode nach Schlaf ein signifikanter Leistungszuwachs gegenüber der Aneignungsphase. Ein durchwachter Behaltenszeitraum führt allerdings zu leichten Leistungsverlusten.

Die Ergebnisse von Rångtjell et al. (2017) zeigen ein ähnliches Bild: Die Ergebnisse unterscheiden sich deutlich, je nachdem welche Methode zur Berechnung der „Learning-Performance“ (Leistung am Ende der Aneignungsphase vs. Bestleistung) herangezogen wird. Hier zeigt sich, dass die besten drei Trials deutlich mehr richtige Sequenzausführungen enthalten als die letzten drei Trials.

Ermüdung am Ende der Aneignungsphase der Peg-Plugging-Task

Pereira et al. (2015) und Rångtjell et al. (2017) nutzten die jeweils besten Aneignungsblöcke zur Leistungsfeststellung in der Aneignungsphase für den Vergleich der Leistungen zwischen Aneignung und Tests. Abbildung 11 zeigt die Aneignungsverläufe der TET für drei Experimente mit dem Pegboard aus Publikation 4. Hieraus wird ersichtlich, dass – über alle VPs gemittelt – die drei jeweils letzten Blöcke der Aneignungsphase fast immer (Ausnahmen: Experiment 2 – Gruppe EM; Experiment 3 – Gruppe ME; siehe Abbildung 11) diejenigen mit der besten Leistung (gemittelt über die VPs einer Gruppe, kürzeste TET = schnellste Sequenzausführungen) waren. Zwar kann dadurch Ermüdung zum Ende der Aneignungsphase nicht per se ausgeschlossen werden, allerdings wird dadurch indiziert, dass Ermüdung zum Ende der Aneignungsphase

keineswegs als ein allgemein aufgabenübergreifender leistungslimitierender Faktor zu veranschlagen ist.

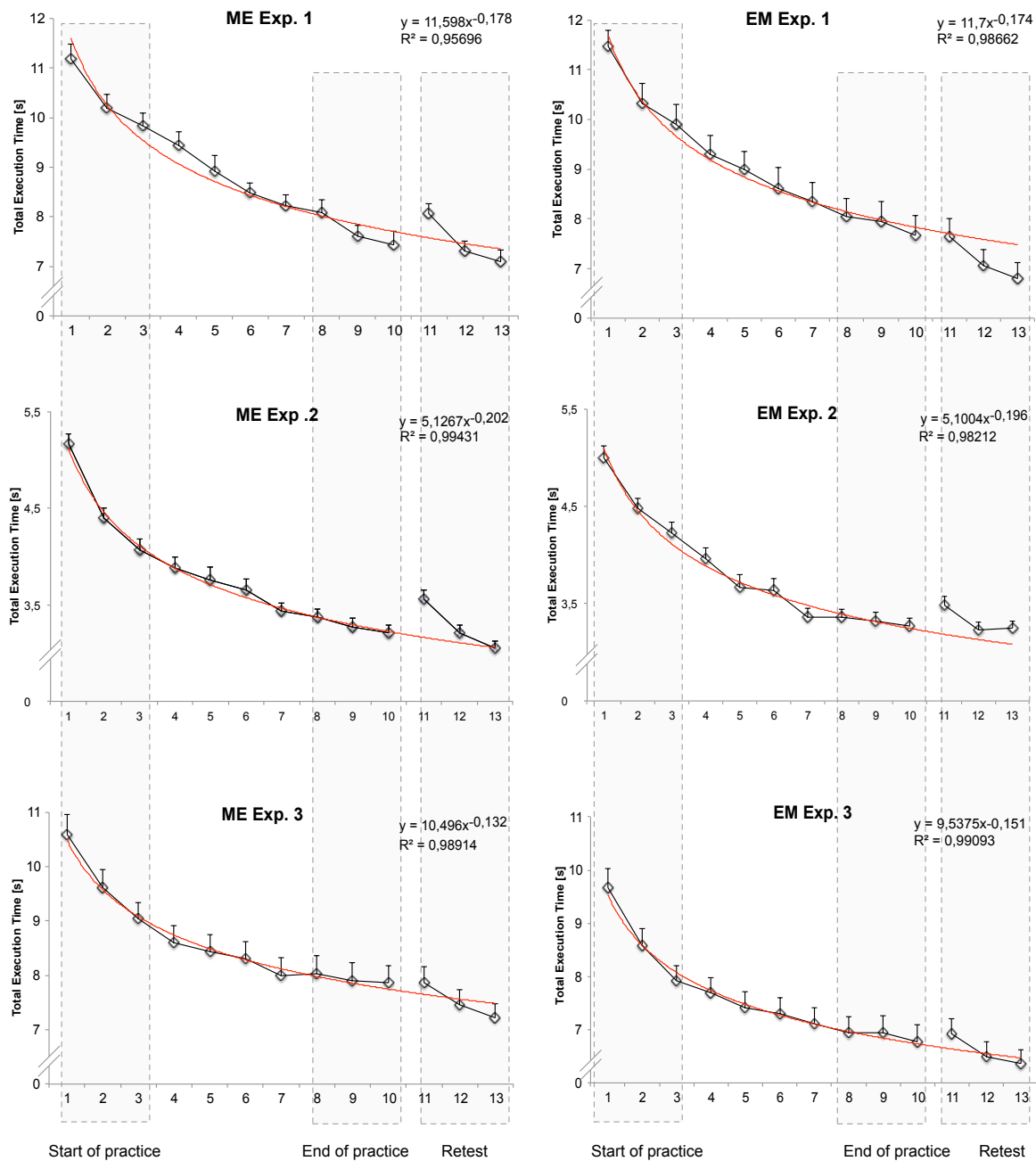


Abbildung 11: Darstellung der Aneignungsverläufe der Total Execution Time aller Experimentalgruppen der drei Experimente aus Publikation 4.

Um dennoch zu überprüfen, ob sich für die Experimente mit dem Peg-Plugging-Paradigma andere Ergebnisse durch die von Pereira et al. (2015) und Rångtjell et al. (2017) vorgeschlagene Auswertung ergeben, wurde mit den Daten aus Publikation 4

zusätzlich eine Auswertung der Change-Scores, einmal unter Berücksichtigung der jeweils letzten drei Aneignungsblöcke im Vergleich zu allen drei Testblöcken sowie einmal unter Berücksichtigung der individuell besten drei (im Vergleich zu allen drei Testblöcken) bzw. besten zwei Aneignungsblöcke (im Vergleich zu den besten zwei Testblöcken), durchgeführt.

Mittels einer 3 „Blockauswahl“⁷ \times 2 „Retentionsbedingung“ [Schlaf – Wachheit] \times 2 „Komplexität“ [hoch⁸ – niedrig] ANOVA wurde ermittelt, ob sich die Wahl der Berechnungsmethode auf das Gesamtergebnis auswirkt. Hierzu wurden die Change-Scores des jeweils ersten 12-stündigen Behaltensintervalls herangezogen – mit dem Ergebnis, dass dies nicht der Fall ist. Zwar gibt es einen signifikanten Haupteffekt „Blockauswahl“ ($F_{[\text{Blockauswahl}]}(2) = 23,236; p < .001$), was bedeutet, dass sich die Ergebnisse der Berechnungsmethoden signifikant voneinander unterscheiden, allerdings gibt es weder einen signifikanten Interaktionseffekt des Faktors „Blockauswahl“ mit einem der anderen Faktoren (Retentionsbedingung, Komplexitätsgrad) noch eine signifikante Dreifachinteraktion, was bedeutet, dass sich die durch den Faktor „Blockauswahl“ definierten Change-Scores nicht signifikant in Abhängigkeit von einem der anderen Faktoren verändern. Das wiederum bedeutet, dass sich die Ergebnisse zum Einfluss von Retentionsbedingung und Komplexitätsgrad im Grundsatz nicht von den ursprünglichen Ergebnissen (drei letzte Aneignungsblöcke) unterscheiden. In Abbildung 12 sind die Change-Scores für die verschiedenen Berechnungsmethoden visuell dargestellt.

⁷ (Change Scores: Veränderung zwischen den drei letzten Aneignungsblöcken und allen drei Testblöcken)
– (Change Scores: Veränderung zwischen den drei besten Aneignungsblöcken und allen drei Testblöcken)
– (Change Scores: Veränderung zwischen den zwei besten Aneignungsblöcken und den zwei besten Testblöcken)

⁸ Hohe Komplexität: Experimente Peg-15 & Peg-10; niedrige Komplexität: Experimente Peg-10reg & Peg-5

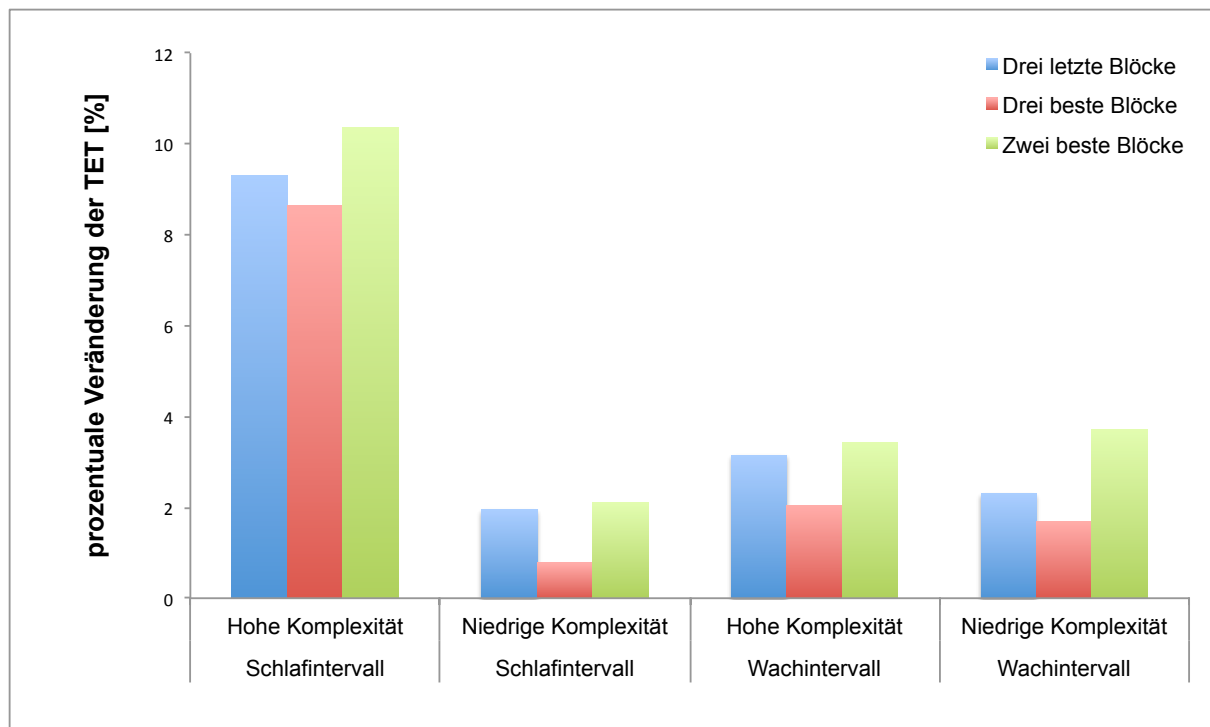


Abbildung 12: Dargestellt ist die prozentuale Leistungsverbesserung für unterschiedliche Datengrundlagen, aufgeteilt in Sequenzen mit hoher Komplexität (Peg-15 & Peg-10) und Sequenzen mit niedriger Komplexität (Peg-10reg & Peg-5) sowie in Wachintervall und Schlafintervall. Eine größere prozentuale Veränderung bedeutet dabei eine stärker verringerte TET und somit eine Leistungsverbesserung.

5.1.6 Mit Übungsende erreichtes Leistungsniveau und schlafbasierte Leistungssteigerungen

Rångtjell et al. (2017) zeigten, dass das mit Übungsende erreichte Leistungsniveau invers mit der Leistung in den Behaltenstests zusammenhängt, und zwar unabhängig von Schlaf oder Wachheit im Behaltensintervall. Probanden mit einer guten Lernleistung (Leistung am Ende der Aneignungsphase) konnten ihre Leistung weniger stark steigern als Probanden mit schwächerer Lernleistung während der Aneignungsphase.

Eine Regressionsanalyse mit Daten der Experimente Peg-15, Peg-10, Peg-10reg sowie Peg-5 (vgl. Publikationen 1, 2 und 4) zeigte, dass ein solcher inverser Zusammenhang für die Experimente mit dem Pegboard nicht besteht. Für die Bedingungen mit hoher Aufgabenkomplexität (Peg-15 & Peg-10) fand sich gar kein Zusammenhang zwischen der Leistung am Ende der Aneignungsphase (gemittelt über die jeweils letzten drei Aneignungsblöcke und die Leistungssteigerung im Retest nach einem 12-stündigen

Behaltensintervall ($R^2 = 0,018$; $F(1, 46) = 0,822$; $p = .369$)). Allerdings fand sich für die Bedingung mit niedriger Aufgabenkomplexität (Peg-5 & Peg-10reg) ein zwar geringer, aber doch signifikanter positiver Zusammenhang zwischen Leistung am Aneignungsende und Veränderung während des Behaltenszeitraumes ($R^2 = 0,251$, $F(1, 47) = 15,717$; $p < .001$). Probanden, die eine gute „Aneignungsperformance“ aufwiesen, konnten ihre Leistung stärker steigern als Probanden mit geringerer Leistung.

Beachtenswert ist in diesem Zusammenhang, dass die Leistungsveränderungen dieser Probanden über den Behaltenszeitraum hinweg keineswegs in ihrer Richtung global übereinstimmten. In der Bedingung mit niedriger Aufgabenkomplexität standen 30 Veränderungen mit positivem Vorzeichen (Leistungssteigerung) 19 Werten mit negativem Vorzeichen gegenüber (Verschlechterung der Leistung im Retest). Somit ergab sich zwar ein leichter positiver Zusammenhang, allerdings im Mittel keine signifikante Leistungssteigerung im Retest. Insgesamt ergab sich kein einheitlicher Zusammenhang zwischen der Aneignungsleistung und der Leistungsveränderung während des Behaltenszeitraumes. Ein inverser Zusammenhang konnte jedoch in keiner der Komplexitätsbedingungen gezeigt werden.

5.2 Fazit: Keine Verzerrung der statistischen Auswertung bei der Peg-Plugging-Aufgabe durch Datenmittelung von Leistungsdaten in den hier vorgelegten Untersuchungen

Es existieren einige Phänomene, die durch Datenmittelung über mehrere Blöcke zu einer Verschleierung der „wahren“ Performance am Ende einer Aneignungsphase führen können und bei der Planung und Auswertung zukünftiger Experimente berücksichtigt werden müssen. Einige Beispiele hierfür sind Ermüdung, reaktive Hemmung, Early Boost sowie Warm-up-Effekte (siehe Rickard et al., 2008; Nettersheim et al., 2015; Hotermans et al., 2006). Allerdings scheint das Auftreten einiger der genannten Phänomene abhängig vom jeweiligen Paradigma und dem Experimentalprotokoll und damit auch von den Anforderungen der jeweiligen Aufgabe an den Organismus zu sein. So konnte oben in den entsprechenden Unterkapiteln gezeigt werden, dass bspw. reaktive Hemmung und Ermüdung im Peg-Plugging-Paradigma während der Aneignungsphase nicht auftreten.

Im Gegensatz dazu sind diese Phänomene beim SRTT-Paradigma wiederholt beobachtet worden.

Weiterer Untersuchung bedarf das transiente Auftreten des Early Boosts. Hier ist noch nicht geklärt, ob der Early Boost durch frühe Konsolidierungsprozesse, durch einen Rückgang von reaktiver Hemmung und Ermüdung oder durch einen anderen Mechanismus zustande kommt. Im Kontext des Peg-Plugging-Paradigmas ist in weiteren Untersuchungen zu klären, ob ein Early Boost überhaupt auftritt. In Publikation 3 konnte gezeigt werden, dass der Early Boost 15 Minuten nach Aneignungsende bei der Peg-Plugging-Aufgabe nicht auftritt. Allerdings konnte dies bislang nur unter Free-Recall-Bedingungen, also unter Abwesenheit von Stimulusinformation, gezeigt werden. Hier gilt es zu untersuchen, ob unter Verfügbarkeit von Stimulusinformation ebenso ein Early Boost auftritt.

Die genannten Phänomene erfordern für zukünftige Untersuchungen eine präzise Planung, eine intensive Auseinandersetzung mit den Aufgabenbedingungen sowie eine Prüfung, ob und in welchem Ausmaß Datenmittelung über mehrere Ausführungsblöcke die wahre Performance verschleiern könnte.

Obwohl die Ergebnisse einiger Autoren (Rickard et al., 2008; Hotermans et al., 2006; Brawn et al., 2010; Pan & Rickard, 2015; Nettersheim et al., 2016) in Bezug auf die oben genannten Phänomene die Existenz schlafbasierter Leistungssteigerungen infrage stellen, geben die Experimente aus den Publikationen 1 bis 4 klare Hinweise auf ebendiese schlafbasierten Leistungssteigerungen. Die meisten Phänomene, die zu einer Überschätzung von Unterschieden zwischen Aneignungsende und Retest führen können, scheinen bei dem Peg-Plugging-Paradigma nicht aufzutreten. Zudem wurde bei der statistischen Auswertung konservativ vorgegangen (bspw. Einbeziehung des Warm-up-Blockes in die Auswertung) und den Empfehlungen der oben angeführten Autoren hinsichtlich der Auswertung gefolgt (Curve-Fitting). Darüber hinaus konnte gezeigt werden, dass schlafbasierte Leistungssteigerungen in der Peg-Plugging-Aufgabe zumindest in Teilen durch Optimierungen der Gedächtnisrepräsentation bzw. entsprechender Retrieval-Strategien zu erklären sind und nicht etwa einer optimierten Stimuluswahrnehmung oder -verarbeitung oder einer verstärkten Stimulus-Response-Assoziationsbildung zugeschrieben werden können (Publikation 3). Aus Publikation 2 wird deutlich, dass qualitative Veränderungen (Chunk-Concatenation) an der Gedächtnisrepräsentation für wesentliche Teile der Leistungssteigerungen

verantwortlich sind. Allerdings unterscheiden sich die Aufgabenanforderungen des Peg-Plugging-Paradigmas in wesentlichen Punkten von denen der Fingertapping-Task.

Insgesamt ist in Betracht zu ziehen, dass die alleinige Erhebung von Leistungsdaten für die detaillierte Untersuchung von Konsolidierungsprozessen nicht ausreichend ist. Bspw. ist es denkbar, dass sich qualitative Veränderungen der Sequenzrepräsentation ebenso in weiteren Parametern (wie bspw. der Blickführung oder kinematischen Daten) zeigen können. Polysomnografische Daten (EEG-, EOG- und EMG-Aufzeichnungen) können zudem ebenso wie bildgebende Verfahren weitere Erkenntnisse über Konsolidierungsprozesse geben (erste Hinweise hierzu bei Nopper et al., 2016).

Für zukünftige Untersuchungen sollten diese Erkenntnisse bei der Wahl des Paradigmas und der Planung des jeweiligen Experimentalprotokolls berücksichtigt werden. Zudem sollte die Frage gestellt werden, welche weiteren nicht unmittelbar kriteriumsleistungsbezogenen Parameter für die Untersuchung von Konsolidierungsmechanismen geeignet sein könnten. Diese Frage scheint allerdings eng mit der Wahl des jeweiligen Paradigmas – und damit der Aufgabenanforderungen – verknüpft und kann daher nicht pauschal beantwortet werden.

Schlussendlich lässt sich für das Thema der methodischen Problematik feststellen, dass bei der traditionellen Erhebung von Daten die wahre Leistungsfähigkeit zum jeweiligen Untersuchungszeitpunkt durch Bildung von Mittelwerten verzerrt werden kann. Da noch keine allgemein anerkannte, standardisierte Vorgehensweise existiert, welche sämtliche genannten Punkte befriedigend adressiert, gilt es bei Untersuchungen schlafbasierter Leistungssteigerungen besonderes Augenmerk auf diese Thematik zu richten. Damit einhergehend ist es angebracht, das verwendete experimentelle Paradigma darauf zu prüfen, ob solche Verzerrungen auftreten können, und ggf. einen Weg zu finden, damit adäquat umzugehen

Generell ergibt sich aus diesen Überlegungen die Frage, ob die alleinige Erhebung von globalen Leistungsmaßen (wie globale Ausführungsdauern und Fehlerzahlen) ausreichend ist, um detaillierte Rückschlüsse auf Konsolidierungsprozesse zu ziehen.

Zahlreiche Studien bedienen sich bereits bildgebender Verfahren oder der Polysomnografie (EEG, EMG, EOG) zur Untersuchung von Konsolidierungsprozessen, um weitere Erkenntnisse zu sammeln. Für das Peg-Plugging-Paradigma wurde von Nopper

et al. (2016) bereits eine erste (Pilot-)Untersuchung unter Einbeziehung von polysomnografischen Daten durchgeführt. Sie fanden eine signifikant erhöhte Schlafdauer nach Aneignung der Peg-Plugging-Aufgabe gegenüber einer Kontrollnacht nach Durchführung einer Kontrollaufgabe. Zudem fanden sich in einer vorläufigen Auswertung Hinweise auf eine Korrelation der Schlafspindeldichte in der Lernhemisphäre mit Offline-Leistungssteigerungen in der Peg-Plugging-Aufgabe.

5.3 Weitere Aspekte für künftige Untersuchungen

In diesem Unterkapitel finden sich weitere Aspekte, die für die weitere Untersuchung schlafbasierter Konsolidierungsprozesse eine Rolle spielen.

5.3.1 Time of Testing

Laut Pan & Rickard (2015) spielt weniger der zirkadiane Rhythmus an sich als vielmehr die verstrichene Zeit seit dem letzten Schlaf eine Rolle für die aktuelle Leistungsfähigkeit. Somit beeinflusst dieser Faktor möglicherweise die Ergebnisse von Untersuchungen, die Schlaf vs. Wachheit als Treatment heranziehen. Allerdings sollte laut Pan & Rickard (2015) die Leistungsfähigkeit kurz nach Schlaf (morgens) und lange nach Schlaf (abends) annähernd auf dem gleichen Niveau liegen, so dass dieser Faktor bei gängigen Experimentalprotokollen zu vernachlässigen sein sollte (Pan & Rickard, 2015). Des Weiteren wurden bei den Experimenten der vorliegenden Arbeit Protokolle genutzt, welche die zirkadianischen Einflüsse kontrollieren. Hierzu wurden jeweils zwei Gruppen genutzt, die zur jeweils entgegengesetzten Tageszeit mit der Datenaufnahme begannen. Es zeigte sich kein systematischer Einfluss der Startzeit auf das Auftreten von schlafbasierten Leistungsverbesserungen. Ähnliche Ergebnisse fanden sich auch in den Experimenten, die das SRTT-Paradigma nutzten (Walker et al., 2002).

Studien, welche Konsolidierungseffekte im Zusammenhang mit Tagschlaf untersuchten, zeigten ähnliche Ergebnisse, obwohl Aneignungs- und Testzeitpunkte wiederum zu unterschiedlichen Tageszeitpunkten stattfanden (Nishida & Walker, 2007; Morita et al., 2012).

Aus einer anwendungsorientierten Perspektive ist die Orientierung an natürlichen Tag-Nacht-Rhythmen für die Untersuchung von schlafbasierten Leistungssteigerungen sogar angebracht. Denn in der Praxis des täglichen Lebens lassen sich diese natürlichen

Rhythmen über längere Zeiträume nicht ohne weitere Folgen manipulieren. Experimentelle Ergebnisse können daher nur dann in die Anwendung in der Alltagspraxis einfließen, wenn sie sich an den natürlichen Gegebenheiten orientieren.

5.3.2 Naps

Eine Möglichkeit, Tageszeiteffekte zu kontrollieren, stellt die Untersuchung von schlafbasierten Leistungssteigerungen anhand kurzer Tagschlafepisoden dar.

Nishida & Walker (2007) sowie Morita et al. (2012) und weitere Gruppen (bspw. Korman et al., 2007; Backhaus & Junghanns, 2006) untersuchten den Einfluss von kurzen Tagschlafphasen (Naps) auf die Konsolidierung neu gelernter motorischer Aufgaben. Die Ergebnisse ähneln denen der Untersuchungen von Nachtschlaf. Auch hier zeigte sich eine signifikante Leistungssteigerung nur dann, wenn im Behaltenszeitraum geschlafen wurde. Zudem zeigten Korman et al. (2007), dass ein 90-minütiger Tagschlaf die Empfänglichkeit für Interferenzeffekte deutlich verringert.

Allerdings nutzten fast alle Untersuchungen das SRTT-Paradigma (Fingertapping- oder Finger-to-Thumb-Opposition-Task), so dass keine Verallgemeinerung der Befunde stattfinden sollte. Ein genauerer Blick zeigt zudem, dass Nap-Studien innerhalb des Zeitraumes stattfinden, in dem die Gedächtnisrepräsentation frühe Konsolidierungsprozesse durchläuft und anfällig für Interferenz ist (siehe Brashers-Krug et al., 1995). Auch das weiter oben diskutierte Phänomen des Early Boosts fällt in diesen Zeitraum. Somit fällt die Interpretation dieser Ergebnisse schwer, denn sie könnten ebenso auf weitere Faktoren als das Vorhandensein einer Nap-Phase zurückzuführen sein. King et al. (2017) diskutieren die Untersuchungen von schlafbasierten Konsolidierungseffekten anhand von Naps kritisch, da damit einige Einschränkungen einhergehen: Einschlafprobleme, kürzere Schlafdauer und die hohe Wahrscheinlichkeit für das Nichterreichen späterer Schlafphasen (REM-Schlaf) – und damit das Fehlen von den natürlichen Nachtschlaf kennzeichnenden Parametern (King et al., 2017).

5.3.3 Type of Task

Die Ergebnisse von Blischke et al. (2008) sowie weiteren Untersuchungen (Walker et al., 2002; Al-Sharman & Siengsukon, 2013, 2014; Genzel et al., 2012) zeigen, dass schlafbasierte Leistungssteigerungen bei einer Reihe verschiedener Aufgabentypen zu

finden sind. Allerdings existieren ebenso Aufgaben, die keinerlei solche Effekte hervorbringen (siehe Blischke et al., 2008). Dies wird auch von King et al. (2017) thematisiert und in Teilen kritisch reflektiert.

Da verschiedene Aufgabenklassen (bspw. kontinuierliche vs. diskrete Bewegungsausführung, rhythmische Bewegungen, Lokomotion, Bewegungsfolgen im Vergleich zu motorischen Adaptationsaufgaben) sich bezüglich ihrer sensomotorischen Kontrollmechanismen unterscheiden, liegt die Vermutung nahe, dass auch die Konsolidierungsprozesse der jeweiligen Gedächtnisrepräsentation von der Art der verwendeten Aufgabe abhängig sind. Bislang liegen allerdings keine Studien vor, die den Zusammenhang zwischen der Aufgabenart und der Konsolidierung der jeweiligen Gedächtnisrepräsentation systematisch untersuchen. Aus der Übersicht der Literatur scheint sich zu ergeben, dass besonders bei sequenziell organisierten Aufgaben mit hinreichender Aufgabenkomplexität, die zumindest partiell deklarativen Zugang erlauben und unter expliziten Bedingungen angeeignet werden, schlafbasierte Leistungssteigerungen auftreten (Walker et al., 2002; Blischke et al., 2008; Genzel et al., 2014; Malangré & Blischke, 2014). Im Gegensatz dazu scheint dieses Phänomen nur teilweise bei impliziten Aneignungsbedingungen, non-deklarativen Gedächtnisinhalten und Aufgaben, bei denen motorische Adaptation gefordert ist, aufzutreten (Blischke et al., 2008; Hödlmoser et al., 2015; King et al., 2017).

Die genannten Aspekte bieten bisher nur grobe Anhaltspunkte für die Einstufung und Vorhersage von schlafbasierten Leistungssteigerungen. Es bleibt weiterhin offen, welche Aufgabenanforderungen notwendig sind und wie verschiedene Anforderungen miteinander in Zusammenhang stehen.

5.3.4 Testing-Effekt

Ein weiteres Untersuchungsfeld ergibt sich aus dem Ergebnis einer Pilotstudie zu Publikation 3. Hier wurde zunächst untersucht, wie sich die Leistung 12 Stunden nach Aneignung der Kriteriumsaufgabe unter Free-Recall-Bedingungen verhält. Allerdings wurde in dieser Untersuchung auf einen Kurzzeit-Behaltenstest direkt im Anschluss an die Aneignungsphase verzichtet, mit dem Ergebnis, dass in der Gruppe „Morgen-Abend“ (Behaltensintervall im Wachzustand) nur zwei der neun Probanden dazu in der Lage waren, die Sequenz aus zehn Elementen korrekt zu reproduzieren. Im Gegensatz dazu waren in der Abend-Morgen-Gruppe (die geschlafen hatte) neun von elf Probanden dazu

in der Lage, die Sequenz korrekt zu reproduzieren. Offensichtlich kam es in der Morgen-Abend-Gruppe während des Wachintervalls zum Vergessen essenzieller Sequenzbestandteile. Die Leistung der Abend-Morgen-Gruppe stabilisierte sich dagegen über den nächtlichen Behaltenszeitraum (die Leistung bezüglich der Ausführungsdauer (TET) im morgendlichen Free-Recall-Test unterschied sich nicht von der am abendlichen Aneignungsende mit verfügbarer Stimulusinformation). Damit scheint einem frühen Behaltenstest unter Free-Recall-Bedingungen ca. 15 Minuten nach Aneignungsende, wie in der anschließenden Hauptstudie zur Anwendung gekommen, eine besondere Bedeutung zuzukommen. Erstens scheint er dem Vergessen von Sequenzteilen entgegenzuwirken und zweitens eine notwendige Voraussetzung für die folgende schlafassoziierte Optimierung der Sequenzrepräsentation zu sein. Die positive Wirkung eines frühen Tests auf den Abruf von Gedächtnisinhalten ist aus anderen Disziplinen bekannt. So fanden Roediger & Karpicke (2006), dass der Abruf von verbalen Gedächtnisinhalten von einem frühen Test profitiert. Gemeinhin wird dieser positive Effekt als Testing-Effekt bezeichnet (siehe auch Carpenter & Kelly, 2012).

Obwohl die Thematik eines frühen Tests bei Fingertapping-Aufgaben (siehe bspw. Nettersheim et al., 2016) sowie einer visuomotorischen Aufgabe (Reis et al., 2015) von den genannten Forschungsgruppen aufgenommen wurde, sind bislang keine Untersuchungen über die Auswirkungen von frühen Behaltenstests unter Free-Recall-Bedingungen erschienen (Antony & Paller, 2018). Dies wäre natürlich insbesondere a) bei großmotorischen Aufgaben und b) bei Aufgaben, die eine ausgeprägte deklarative Sequenzerinnerungskomponente aufweisen, wünschenswert. Hierzu liegen nach Kenntnis der Autoren bisher weder systematische Analyseversuche noch formulierte Hypothesen vor. Da solch ein früher Behaltenstest allerdings Auswirkungen auf die nachfolgenden Konsolidierungsprozesse zu haben scheint, ergibt sich hier eine Forschungsfrage für künftige Untersuchungen.

5.3.5 Alter und schlafbasierte Leistungsveränderung

Da sich verschiedene Schlafparameter mit zunehmendem Alter verändern (Landolt et al., 1996) und Untersuchungen zu schlafbasierten Leistungssteigerungen bislang vornehmlich an Stichproben mit jungen Erwachsenen vorgenommen wurden, stellt sich die Frage, ob sich die gefundenen Ergebnisse auch auf ältere Personengruppen übertragen lassen. Erste Ergebnisse mit Stichproben unterschiedlichen Alters (Al-

Sharman & Siengsukon, 2013, 2014) unter Verwendung einer großmotorischen Aufgabe zeigen, dass schlafbasierte Leistungsveränderungen in verschiedenen Altersklassen auftreten. So fanden Al-Sharman & Siengsukon (2013, 2014), dass sich nur bei Probanden, die zwischen Aneignungsende und Tests schlafen konnten, die Leistung nach einer Periode der Inaktivität in einer funktionalen Gangaufgabe verbesserte; und dies sowohl bei jungen Erwachsenen als auch bei Probanden mittleren Alters (im Mittel 48 Jahre alt) sowie bei alten Probanden (im Mittel 70 Jahre alt). Sie verwendeten dazu die Verbindung einer Gangaufgabe auf einem für die Probanden neuen, ca. 30 Meter langen, in der konkreten Streckenführung unregelmäßig gestalteten, elliptischen Pfad (gefordert war hier also eine fixierte Sequenz spezifischer Richtungsänderungen) – inklusive einer ca. 3 Meter langen Strecke im Tandemschritt – mit einer kognitiven Zweitaufgabe (Rückwärtszählen in 7er-Schritten von einer randomisiert ausgewählten Zahl zwischen 299 und 293), um erstmals eine funktionale motorische Aufgabe auf schlafbasierte Offline-Learning-Effekte zu untersuchen.

Allerdings ist die Befundlage keinesfalls eindeutig. So fanden bspw. Spencer et al. (2007) anhand zweier modifizierter, 10 Elemente umfassender Versionen der SRTT – einmal unter expliziten Aneignungsbedingungen (die Probanden wurden über die zu lernende Sequenz unterrichtet und der Ort des visuellen Stimulus stimmte mit dem Ort der jeweils zu drückenden Taste überein), einmal unter impliziten Bedingungen (die Probanden hatten keine Bewusstheit über eine zu lernende Sequenz und die zu drückende Taste wurde durch die Farbe des visuellen Stimulus codiert) –, dass schlafbasierte Leistungssteigerungen unter beiden Bedingungen in der SRTT zwar bei jungen Erwachsenen (mittleres Alter: 20,8 Jahre) auftraten, im Gegenzug jedoch keine Verbesserung der Leistung bei einer Gruppe älterer Erwachsener (mittleres Alter: 59 Jahre) gefunden werden konnte.

Ähnliche Ergebnisse berichten Brown et al. (2009) für eine weitere Variante der SRTT mit einem 24-stündigen Behaltensintervall. Hier wurden Sequenzen, welche 12 Elemente umfassten, verwendet und mit randomisierten, nicht der zu lernenden Sequenz entsprechenden Tastendrücken im Wechsel präsentiert. Es erfolgte keine Information darüber, dass die Aufgabe eine Sequenz enthielt. Hier waren die mittelalten Erwachsenen (Durchschnittsalter: 58,3 Jahre) zwar in der Lage, die Aufgabe auf dem Niveau der jüngeren Erwachsenen (Durchschnittsalter: 20,4 Jahre) auszuführen, im Gegensatz zu der

Gruppe junger Erwachsener konnten bei den Älteren jedoch keine Offline-Leistungszuwächse festgestellt werden.

Dorfberger et al. (2007) konnten mit einer Finger-to-Thumb-Opposition-Aufgabe, einer Variante der SRTT, bei der vier Finger einer Hand in einer festgelegten Abfolge in Opposition zum Daumen gebracht werden müssen, zeigen, dass auch Kinder und Jugendliche im Alter von neun, 12 und 17 Jahren zwar ein unterschiedliches Ausgangsniveau in der Kriteriumsleistung aufweisen. Verzögerte Leistungsgewinne 24 Stunden nach der Aneignungsphase traten dennoch in allen drei Altersstufen auf.

Gudberg et al. (2015) schlagen aufgrund ihrer Ergebnisse vor, dass diese uneindeutige Befundlage durch die generell verringerte Fingerfertigkeit bei älteren Personen ("There is evidence to suggest significant reductions in fine motor skill, including speed, dexterity, and finger strength with older age." (Gudberg et al., p. 1410)) zustande kommen mag. In ihrer Untersuchung nahmen sie eine Doppeldissoziation von Alter und Art der Aufgabe vor. Sie nutzten einmal die klassische SRTT mit fünf Elementen und übertrugen diese für die andere (keine feinmotorischen Fingerbewegungen erfordernde) Bedingung in einen größeren Rahmen, in dem unrestringierte Armbewegungen für das Erreichen der entsprechenden Zielpunkte mit der ganzen Hand notwendig waren. Die Ergebnisse zeigen, wie bei den weiter oben genannten Studien auch, dass unter Verwendung der klassischen SRTT bei jungen Erwachsenen (18–35 Jahre) nach Schlaf deutliche Offline-Leistungssteigerungen auftreten. Bei älteren Erwachsenen (50–85 Jahre) zeigte sich dagegen keine Leistungssteigerung nach Schlaf. Umgekehrt verhielt es dagegen bei der Bedingung, die eine großmotorische Bewegungsproduktion erforderte: Hier zeigten sich Offline-Leistungssteigerungen nach einem Schlaf enthaltenden Behaltensintervall nur bei der Gruppe älterer Erwachsener. In der Gruppe der jungen Erwachsenen waren demgegenüber unter dieser Bedingung Leistungssteigerungen unabhängig von Wachheit oder Schlaf zu verzeichnen.

Bottary et al. (2016) nutzten eine 7-teilige Fingerbewegungssequenz mit einer Gruppe junger Erwachsener und einer Gruppe älterer Erwachsener. Hier zeigten sich signifikante Leistungssteigerungen nur in der Gruppe der jungen Erwachsenen. Zudem konnten die Autoren zeigen, dass bei jungen Erwachsenen nur nach dem Schlafintervall Chunk-Concatenation zu beobachten war. Im Gegensatz dazu war dies bei der Gruppe älterer Erwachsener nicht der Fall.

Die uneinheitliche Befundlage lässt bislang nur wenige Schlüsse bezüglich der Auswirkungen des Alters auf das Auftreten von leistungssteigernden schlafbasierten Konsolidierungseffekten zu. Es scheint, dass in Abhängigkeit von dem jeweils genutzten experimentellen Paradigma das Alter der Probanden unterschiedlichen Einfluss auf die Ergebnisse hat. Wie sich die Offline-Leistungssteigerungen bei der in dieser Arbeit verwendeten Peg-Plugging-Aufgabe bei unterschiedlichen Altersgruppen verhalten, kann bis dato nicht gesagt werden, da bislang lediglich jüngere Erwachsene als Probanden untersucht wurden. Zwar handelt es sich um eine Aufgabe, bei der, ähnlich wie bei der Handtapping-Aufgabe von Gudberg et al. (2015), großmotorische Bewegungsproduktion und Redundanzkontrolle eine Rolle spielen, allerdings ist sicherlich auch eine feinmotorische Komponente in der Peg-Plugging-Aufgabe enthalten, da die Bewegungsziele nicht mit der ganzen Handfläche, sondern mit einem kleinen Holzstab erreicht werden müssen, der in ein nur geringfügig größeres Loch gesteckt werden muss. Ebenso wurden in keiner der in dieser Arbeit verwendeten Experimente ähnliche Ergebnisse gefunden, wie sie Gudberg et al. (2015) für die Gruppe junger Erwachsener fanden. Durch diese Befundlage wird deutlich, dass das Alter in Wechselwirkung mit der Aufgabenkomplexität und den effektorspezifischen Kontrollanforderungen ein weiterer Einflussfaktor für das Auftreten leistungssteigernder schlafassoziierter Konsolidierungseffekte sein könnte. Welche Rolle das Alter dabei genau spielt, muss in künftigen Untersuchungen geklärt werden.

5.3.6 Weitere Komplexitätsfaktoren

In Publikation 4 dieser Arbeit wurde erstmals systematisch untersucht, welchen Einfluss die Aufgabenkomplexität auf schlafbasierte Leistungssteigerungen hat. Die Ergebnisse legen nahe, dass eine gewisse Aufgabenkomplexität (operationalisiert über die Sequenzlänge und die strukturelle Komplexität) notwendig ist, damit schlafbasierte Leistungssteigerungen überhaupt beobachtbar werden. Ähnliche Befunde für die SRTT finden sich bei Kuriyama et al. (2012). Allerdings beinhalteten bisherige Untersuchungen lediglich die Veränderung der Effektorzahl, der Sequenzlänge sowie der strukturellen Komplexität der Aufgabe. Es existieren weitere Faktoren, wie die Einbindung von Effektoren beider Körperseiten, veränderte Aufgabenschwierigkeit (Präzisionsanforderungen) oder die Notwendigkeit der Verarbeitung weiteren sensorischen Inputs, welche die Aufgabenkomplexität (Aufwand des hierarchisch

aufgebauten Bewegungskontrollsystems) beeinflussen können. Diese sind, zumindest nach Kenntnis des Autors, bislang nicht systematisch untersucht worden.

Da die Ergebnisse von Publikation 4 zeigen, dass schlafbasierte Leistungssteigerungen bei Ausführung von Armbewegungsfolgen nur bei hinreichender Aufgabenkomplexität auftreten, besteht eine Aufgabe künftiger Untersuchungen unter anderem darin, zu klären, welchen Einfluss weitere Komplexitätsfaktoren auf das Auftreten schlafbasierter Leistungssteigerungen haben. Von besonderem Interesse ist dabei, inwiefern sich die Kombination verschiedener Komplexitätsfaktoren auswirkt. Bspw. zeigen Daten von Schmitt et al. (2010), dass eine Übertragung der SRTT auf ein Smartboard mit sehr niedrigem Fitts' ID keinerlei schlafbasierte Leistungssteigerungen auslöst, unabhängig davon, ob die Aufgabe mit einem oder mehreren Endeffektoren ausgeführt wird. Obwohl mehrere Endeffektoren beider Körperseiten besser koordiniert werden mussten (Anzeichen für hohe Komplexität), waren die Präzisionsanforderungen und die Sequenzlänge vergleichsweise gering. Im Gegensatz dazu untersuchten Kempler & Richmond (2012) eine Aufgabe, die auf den ersten Blick ähnlich geringe Präzisionsanforderungen aufweist (bilaterale Armbewegungssequenz). Allerdings könnte in dieser Untersuchung die für das Auslösen von schlafbedingten Leistungssteigerungen notwendige Komplexität auf anderem Weg erreicht worden sein. Die genutzte Sequenz bestand aus sechs Sequenzelementen. Jedes Sequenzelement bestand dabei aus einer Kombination aus der jeweils richtigen Position beider Arme. Für jeden Arm existierten drei mögliche Positionen: zur Decke zeigend, zur Seite zeigend und auf den Boden zeigend. Es galt also nicht nur, die Bewegungen beider Arme zu koordinieren, sondern es musste auch für jedes der sechs Sequenzelemente eine Auswahlentscheidung zwischen neun möglichen Armpositionskombinationen getroffen werden. Möglicherweise resultierte die für das Auftreten von schlafbedingten Leistungssteigerungen nötige Komplexität in diesem Experiment aus der Anzahl an für die richtige Aufgabenausführung notwendigen Auswahlentscheidungen bzw. Auswahlmöglichkeiten.

Über die genauere Betrachtung verschiedener Faktoren, welche die Aufgabenkomplexität beeinflussen, bietet sich die Chance, weitere Einsicht in die Einflussfaktoren für das Auftreten schlafbasierter Leistungssteigerungen zu erhalten. Damit wäre ein weiterer Schritt in Richtung der Anwendung der Erkenntnisse in Alltagssituationen, wie

sportlichen Bewegungsaufgaben, Rehabilitation oder dem Erlernen von Arbeitsabläufen, getan.

6 Fazit

In diesem Dissertationsvorhaben wurde erstmals systematisch untersucht, ob schlafbasierte Leistungssteigerungen auch bei einer komplexen großmotorischen Armbewegungssequenz auftreten. Es konnte gezeigt werden, dass schlafbasierte Leistungssteigerungen bei der Pegboard Aufgabe unter bestimmten Bedingungen auftreten. Zudem konnte gezeigt werden, dass die Ergebnisse, zumindest in Teilen, auf Optimierungen an der Gedächtnisrepräsentation – in Gestalt einer effizienteren Verknüpfung aufeinanderfolgender Sequenzbausteine (Chunks) – während Konsolidierungsphasen zurückzuführen sind. Als modulierender Faktor für das Auftreten der berichteten schlafbasierten Leistungssteigerungen konnte die Komplexität der verwendeten Aufgabe identifiziert werden. Hier zeigte sich, dass sowohl die Manipulation der Sequenzlänge als auch die strukturelle Komplexität Einfluss auf schlafbasierte Leistungssteigerungen haben.

Weiterhin wurde deutlich, dass methodische Probleme bei der Datenanalyse von seriellen Fingerbewegungssequenzen auftreten können. Diese Probleme scheinen allerdings durch die Charakteristik und Handhabung der Peg-Plugging-Aufgabe im für diese Arbeit verwendeten experimentellen Paradigma hinreichend kontrolliert zu sein. Für zukünftige Untersuchungen scheint es bei der Analyse von Leistungsdaten dennoch sinnvoll, Verfahren zu nutzen, die eine Verzerrung von Datenmittelung minimieren.

Zudem mögen sich die Auswirkungen von schlafbegleitenden Konsolidierungsprozessen auch in weiteren Parametern außerhalb reiner Leistungsdaten zeigen. Hier gilt es für zukünftige Forschungsvorhaben, diese Parameter zu identifizieren und einer Prüfung zu unterziehen.

7 Literaturverzeichnis

Abrahamse, E.L., Ruitenberg, M.F.L., de Kleine, E. & Verwey, W.B. (2013). Control of automated behavior: insights from the discrete sequence production task. *Frontiers in Human Neuroscience*, 7. doi: 10.3389/fnhum.2013.00082

Ajemian, R., Dausilio, A., Moorman, H. & Bizzi, E. (2010). Why Professional Athletes Need a Prolonged Period of Warm-Up and Other Peculiarities of Human Motor Learning. *Journal of Motor Behaviour*, 42(6), 381 – 388. doi:10.1080/00222895.2010.528262

Al-Sharman, A. & Siengsukon, C.F. (2013). Sleep Enhances Learning of a Functional Motor Task in Young Adults. *Physical Therapy*, 93 (12), 1625-1635. doi: 10.2522/ptj.20120502

Al-Sharman, A. & Siengsukon, C.F. (2014). Performance on a Functional Motor Task Is Enhanced by Sleep in Middle-Aged and Older Adults. *Journal of Neurologic Physical Therapy*, 38, 161 – 169. doi: 10.1097/NPT.0000000000000048

Antony, J.W. & Paller, K.A. (2018). Retrieval and sleep both counteract the forgetting of spatial information. *Learning & Memory*, 25, 258-263. doi: 10.1101/lm.046268.117

Appleman, E.R., Albouy, G., Doyon, J., Cronin-Golomb & A., King, B.R. (2016). Sleep Quality Influences Subsequent Motor Skill Acquisition. *Behavioral Neurosciences*, 130 (3), 290–297. doi: 10.1037/bne0000131

Backhaus, J. & Junghanns, K., (2006). Daytime naps improve procedural motor memory. *Sleep Medicine*, 7 (6), 508–512. doi: 10.1016/j.sleep.2006.04.002

Bizzi, E. & Ajemian, R. (2015). A Hard Scientific Quest: Understanding Voluntary Movements. *MIT Press*, 144(1), 83-95. doi: [10.1162/DAED_a_00324](https://doi.org/10.1162/DAED_a_00324)

Blischke, K., Erlacher, D., Kresin, H., Brückner, S. & Malangré, A. (2008). Benefits of Sleep in Motor Learning - Prospects and Limitations. *Journal of Human Kinetics*, 20 (1), 23 – 35. doi: 10.2478/v10078-008-0015-9

Blischke, K. & Malangré, A. (2016). Chunk Concatenation Evolves with Practice and Sleep-Related Enhancement Consolidation in a Complex Arm Movement Sequence. *Journal of Human Kinetics*, 51 (1), 5- 17. doi: 10.1515/hukin-2015-0163.

Blischke, K. & Malangré, A. (2017). Task Complexity Modulates Sleep-Related Offline Learning in Sequential Motor Skills. *Frontiers in Human Neuroscience*, 11:374. doi: 10.3389/fnhum.2017.00374

Blischke, K. & Müller, H. (2009). Motorisches Lernen. In: Schlicht, W. & Strauß, B. (Hrsg.) *Grundlagen der Sportpsychologie*. Sportpsychologie Band 1. (159-228). Hogrefe, Göttingen.

Brashers-Krug, T., Shadmehr, R. & Bizzi, E. (1996). Consolidation in human motor memory. *Nature*, 382, 252 – 255.

Brawn, T.P., Fenn, K.M., Nusbaum, H.C. & Margoliash, D. (2010). Consolidating the effects of waking and sleep on motor-sequence learning. *The Journal of Neuroscience*, 30(42), 13977 -13982. Doi: 10.1523/JNEUROSCI.3295-10.2010

Brown, R.M., Robertson, E.M. & Press D.Z. (2009). Sequence skill acquisition and off-line learning in normal aging. *PLos One*, 4(8). doi: 10.1371/journal.pone.0006683

Born, J. & Wilhelm, I. (2012). System consolidation of memory during sleep. *Psychological Research*, 76, 192 – 203. doi: 10.1007/s00426-011-0335-6

Bottary, R., Sonni, A., Wright, D. & Spencer, R.M.C. (2016). Insufficient chunk concatenation may underlie changes in sleep-dependent consolidation of motor sequence learning in older adults. *Learning & Memory*, 23, 455- 459. doi: 10.1101/lm.043042.116

Boyce, W.F., Gowland, C., Rosenbaum, P.L., Lane, M., Plews, N., Glodsmith, C.H., Russell, D.J., Wright, V., Potter, S. & Harding, D. (1995). The Gross Motor Performance Measure: Validity and Responsiveness of a Measure of Quality of Movement. *Physical Therapy*, 75(7), 603 – 613.

Carpenter, S.K. & Kelly, J.W. (2012). Tests enhance retention and transfer of spatial learning. *Psychonomic Bulletin & Review*, 19, 443-448. doi: 10.3758/s13423-012-0221-2

Christova, M., Aftenberger, H., Nardone, R. & Gallasch, E. (2018). Adult Gross Motor Learning and Sleep: Is There a Mutual Benefit? *Neural Plasticity* 2018. doi: 10.1155/2018/3076986

Dahl, A. (1928). Über den Einfluß des Schlafens auf das Wiedererkennen. *Psychologische Forschung* 11(1), 290-301.

Dayan, E. & Cohen, L.G. (2011). Neuroplasticity Subservient Motor Skill Learning. *Neuron*, 72, 443 – 454. doi: 10.1016/j.neuron.2011.10.008

De Vivo, L., Bellesi, M., Marshall, W., Bushong, E.A., Ellisman, M.H., Tononi, G. & Cirelli, C. (2017). Ultrastructural evidence for synaptic scaling across the wake/sleep cycle. *Science*, 355 (6324), 507-510. doi: 10.1126/science.aah5982

Diedrichsen, B. & Kornysheva, K. (2015). Motor skill learning between selection and execution. *Trends in Cognitive Science*, 19(4), 227-233. doi: 10.1016/j.tics.2015.02.003

Dorfberger, S., Adi-Japha, E. & Karni, A. (2007). Reduced Susceptibility to interference in the consolidation of motor memory before adolescence. *PLoS ONE*, 2(2): e240. doi:10.1371/journal.pone.0000240

Dorfberger, S., Adi-Japha, E. & Karni, A. (2012). Sequence Specific Motor Performance Gains after Memory Consolidation in Children and Adolescents. *Plos ONE*, 7(19): e28673. doi:10.1371/journal.pone.0028673

Doyon, J. & Benali, H. (2005). Reorganization and plasticity in the adult brain during learning of motor skills. *Current Opinion in Biology*, 15, 161-167. doi: 10.1016/j.conb.2005.03.004

Fischer, S., Hallschmidt, M., Elsner, A. L., & Born, J. (2002). Sleep forms memory for finger skills. *Proceedings of the National Academy of Sciences*, 99, 11987-11991.

Fischer, S., Nitschke, M.F., Melchert, U.H., Erdmann, C. & Born, J. (2005). Motor Memory Consolidation in Sleep Shapes More Effective Neuronal Representations. *Journal of Neuroscience*, 25 (49), 11248-11255. doi: 10.1523/JNEUROSCI.1743-05.2005

Fogel, S., Vien, C., Karni, A., Benali, H., Carrier, J. & Doyon, J. (2017). Sleep spindles: a physiological marker of age-related changes in gray matter in brain regions supporting motor skill memory consolidation. *Neurobiology of Aging*, 49. doi: 10.1016/j.neurobiolaging.2016.10.009

Gais, S., Lucas, B. & Born, J. (2006). Sleep after learning aids memory recall. *Learning & Memory*, 13, 259-262. doi: 10.1101/lm.132106

Genzel, L., Quack, A., Jäger, E., Konrad, B., Steiger, A. & Dresler, M. (2012). Complex Motor Sequence Skills Profit from Sleep. *Neuropsychobiology*, 66 (4), 237 - 243. doi:10.1159/000341878

Genzel, L., Kroes, M.C.W., Dresler, M. & Battaglia, F.P. (2014). Light sleep versus slow wave sleep in memory consolidation: a question of global versus local processes? *Trends in Neuroscience*, 37(1), 10 - 19. doi: 10.1016/j.tins.2013.10.002

Genzel, L. & Robertson, E.M. (2015). To Replay, Perchance to Consolidate: *PLoS Biology*, 13(10): e1002285. doi:10.1371/journal.pbio.1002285

Gudberg C., Wulff, K. & Johansen-Erg, H. (2015). Sleep-dependent motor memory consolidation in older adults depends on task demands. *Neurobiology of Aging*, 36 (3), 1409-1416. doi: 10.1016/j.neurobiolaging.2014.12.014

Herrmann, C., Gerlach, E. & Seelig, H. (2015). Motorische Basiskompetenzen in der Grundschule. Begründung, Erfassung und empirische Überprüfung eines Messinstrumentes. *Sportwissenschaft*. doi: 1.1007/s12662-015-0378-8.

Hotermans, C., Peigneux, P., de Noordhout, A.M., Moonen, G. & Maquet, P. (2006). Early boost and slow consolidation in motor skill learning. *Learning & Memory*, 13, 580-583. doi: 10.1101/lm.239406

Hödlmoser, K. Birklbauer, J., Schabus, M., Eibenberger, P., Rigler, S. & Muelser, E. (2015). The impact of diurnal sleep on the consolidation of a complex gross motor adaptation task. *Journal of Sleep Research*, 24 (1), 100-109. doi: 10.1111/jsr.12207

Hupbach, A. (2018). Longterm effects of directed forgetting. *Memory*, 26 (3), 321 – 329. doi:10.1080/09658211.2017.1358748

Joseph, M.E., King, A.C. & Newell, K.M. (2013). Task Difficulty and the Time Scales of Warm-Up and Motor Learning. *Journal of Motor Behavior*, 45 (3), 231 – 238. doi: 10.1080/00222895.2013.784240

Kempler, L & Richmond, J.L. (2012). Effects of sleep on gross motor memory. *Memory*, 20 (8), 907-914. doi:10.1080/09658211.2012.711837.

King, B.R., Hoedlmoser, K., Hirschauer, F., Dolfen, N. & Albouy, G. (2017). Sleeping in the motor engram: The multifaceted nature of sleep related motor memory consolidation. *Neuroscience and Biobehavioral Reviews*, 80, 1-22. doi: 10.1016/j.neubiorev.2017.04.026

Korman, M., Doyon, J., Doljansky, J., Carrier, J., Dagan, Y. & Karni, A. (2007). Daytime sleep condenses the time course of motor memory consolidation. *Nature Neuroscience*, 10 (9), 1206 – 1231. doi: 10.1038/nn1959

Kuriyama, K., Stickgold, R. & Walker, M.P. (2004). Sleep-dependent learning and motor-skill complexity. *Learning & Memory*, 11, 705-713. doi:10.1101/lm.76304

Landolt, H., Dijk, D., Achermann, P. & Borbély, A. A. (1996). Effect of age on the sleep EEG: slow-wave activity and spindle frequency activity in young and middle-aged men. *Brain Research*, 738 (29), 205-212. doi: 10.1016/S0006-8993(96)00770-6

Malangré, A., Leinen, P. & Blischke, K. (2014). Sleep-related offline learning in a complex

arm movement sequence. *Journal of Human Kinetics*, 40 (1), 7 – 20. doi: 10-2478/hukin-2014-0002

Malangré, A. & Blischke, K. (2016). Sleep-Related Offline Improvements in Gross Motor Task Performance Occur Under Free Recall Requirements. *Frontiers in Human Neuroscience*, 10. doi: 10.3389/fnhum.2016.00134

McGaugh, J.L. (2000). Memory – a Century of Consolidation. *Science*, 287, 248 - 251. doi: 10.1126/science.287.5451.248

Morita, Y., Ogawa, K. & Uchida, S. (2012). The effect of a daytime 2-hour nap on complex motor skill learning. *Sleep and Biological Rhythms*, 10, 302-309. doi:10.1111/j.1479-8425.2012.00576.x

Nacson, J. & Schmidt, R.A. (1971). The activity-set hypothesis for warm-up decrement. *Journal of Motor Behavior*, 3(1), 1-15. doi: 10.1080/00222895.1971.10734887

Nettersheim, A., Hallschmid, M., Born, J. & Diekelmann, S. (2015). The Role of Sleep in Motor Sequence Consolidation: Stabilization Rather Than Enhancement. *The Journal of Neuroscience*, 35 (17), 6696-6702. doi: 10.1523/JNEUROSCI.1236-14.2015

Nishida, M. & Walker, M.P. (2007). Daytime Naps, Motor Memory Consolidation and Regionally Specific Sleep Spindles. *PLoS ONE*, 2 (4), e341. doi:10.1371/journal.pone.0000341

Nopper, I., Krewer, C., Weber, F., Brunner, H., & Blischke, K. (2016). Sleep spindle density and gross motor sequence learning. *Journal of Sport and Exercise Psychology*, 38, (Supplement), 91-92.

Nopper, I., Krewer, C., Weber, F., Brunner, H., & Blischke, K. (2016). Gross motor sequence learning and sleep spindle density. In *International Conference on Sleep Spindling – Book of Abstracts/Conference Programme* (p. 58). Budapest: Akadémiai Kiadó.

Patrick, G.T.W. & Gilbert, J.A. (1896). Studies from the psychological laboratory of the University of Iowa: On the effects of loss of sleep. *Psychological Review*, 3(5), 469-483. doi: 10.1037/h0075739

Pan, S.C. & Rickard, T.C. (2015). Sleep and Motor Learning: Is There Room for Consolidation? *Psychological Bulletin*, 141(4), 812-834. doi: 10.1037/bul0000009

Pereira, S., Beijamini, F. & Louzada, F.M. (2015). Re-examining sleep's effect on motor skills: How to access performance on the finger tapping task? *Sleep Science*. doi:

10.1016/j.slsoci.2015.01.001

Pilcher, J.J. & Huffcutt, A.I. (1996). Effects of Sleep Deprivation on Performance: A Meta Analysis. *Sleep*, 19 (4) 318-326. doi: 10.1093/sleep/19.4.318

Plihal, W. & Born, J. (1997). Effects of Early and Late Nocturnal Sleep on Declarative and Procedural Memory. *Journal of Cognitive Neuroscience*, 9(4), 534-547. doi: 10.1162/jocn.1997.9.4.534

Ramanathan, D.S., Gulati, T., Ganguly, K. (2015). Sleep-Dependent Reactivation of Ensembles in Motor Cortex Promotes Skill Consolidation. *PLoS Biology*, 13 (9). doi: 10.1371/journal.pbio.1002263

Rångtell, F.H., Karamchedu, S., Andersson, P., van Egmond, L., Hultgren, T., Broman, J-E., Cedernas, J. & Benedikt, C. (2017). Learning performance is linked to memory consolidation across both sleep and wakefulness. *Scientific Reports*, 7. doi: 10.1038/s41598-017-09263-5

Reis, J., Fischer, J.T., Prichard, G., Weiller, C. Cohen, L.G. & Fritsch, B. (2015). Time- but not sleep- dependent Consolidation of tDCS-Enhanced Visuomotor Skills. *Cereb Cortex*, 25 (1), 109-117. doi: 10.1093/cercor/bht208

Rickard, T.C., Cai, D.J., Rieth, C.A., Hones, J. & Ard, M.C. (2008). Sleep Does Not Enhance Motor Sequence Learning. *Journal of Experimental Psychology*, 34 (4), 834 -842. doi: 10.1037/0278-7393.34.4.834

Roberston, E.M. & Cohen, D.A. (2006). Understanding Consolidation through the Architecture of Memories. *The Neuroscientist*, 12 (3), 261-271. doi: 10.1177/1073858406287935

Roediger, III, H.L. & Karpicke, J.D. (2006). The Power of Testing Memory. Basic Research and Implications for Educational Practice. *Perspective on Social Science*, 1 (3), 181-210. doi: 10.1111/j.1745-6916.2006.00012.x

Russel, D.J., Rosenbaum, P.L., Cadman, D.T., Gowland, C., Hardy, S. & Jarvis, S. (1989). The gross motor function measure: a means to evaluate the effects of physical therapy. *Developmental Medicine and Child Neurology*, 31, 341-352. doi: 10.1111/j.1469-8749.1989.tb04003.x

Sakai, K., Kitaguchi, K. & Hikosaka, O. (2003). Chunking during human visuomotor sequence learning. *Experimental Brain Research*, 152 (2), 229-242. doi: 10.1007/s00221-003-1548-8

Schaal, S., Sternad, D. Osu, R. & Kawato, M. (2004). Rhythmic arm movement is not discrete. *Nature Neuroscience*, 7, 1136-1143.

Schmidt, R.A. & Wrisberg, C.A. (1971). The activity-set hypothesis for warm up decrement in a movement-speed task. *Journal of Motor Behavior*, 3(4), 318 – 325. doi: 10.1080/00222895.1971.10734912

Schmidt, S., Erlacher, D., Blischke, K., Brueckner, S. & Müller, F. (2010). Sleep related enhancement in motor performance: gross vs. fine motor tasks. *Journal of Sport and Exercise Psychology* 32 (Suppl): 124.

Schmitz, L., Waßmuth, N. (2013). Zum Einfluß der Übungsrate auf Ausführungsgeschwindigkeit und Reproduktionssicherheit einer komplexen Armbewegungsfolge – Eine Laborstudie. *Unpublished Bachelors' Thesis*; Saarland University, Department of Sport Science.

Shadmehr, R. & Holcomb, H.H. (1999). Inhibitory control of competing motor memories. *Experimental Brain Research*, 126 (2), 235 -251. doi: 10.1007/s002210050733

Spencer, R.M.C., Goouw, A.M. & Ivry, R.B. (2007). Age-related decline of sleep-dependent consolidation. *Learning & Memory*, 14, 480-484. doi: 10.1101/lm.569407

Tilley, A.J. & Empson, J.A.C. (1978). REM sleep and memory consolidation. *Biological Psychology*, 6(4), 293–300. doi: 10.1016/0301-0511(78)90031-5

Tucker, M., McKinley, S. & Stickgold, R. (2011). Sleep Optimizes Skill in the Elderly. *Journal of the American Geriatrics Society*, 59(4), 603-609. doi: 10.1111/j.1532-5415.2011.03324.x

Voelcke-Rehage, C. (2008). Motor-skill learning in older adults – a review of studies on age-related differences. *European Review of Aging and Physical Activity*, 5 (30). doi:10.1007/s11556-008-0030-9

Verwey W.B. & Eikelboom, T. (2003). Evidence for Lasting Sequence Segmentation in the Discrete Sequence-Production Task. *Journal of Motor Behavior*, 35 (2), 171-181. doi: 10.1080/00222890309602131

Verwey, W.B., Shea, C.H. & Wright, D.L. (2015). A cognitive framework for explaining serial processing and sequence execution strategies. *Psychonomic Bulletin & Review*, 22 (1), 54-77. doi: 10.3758/s13423-014-0773-4

Walker, M.P., Brakefield, T., Morgan, A., Hobson, J.A. & Stickgold, R. (2002). Practice with

Sleep Makes Perfect: Sleep-Dependent Motor Skill Learning. *Neuron*, 35 (1), 205 – 211. doi: 10.1016/S0896-6273(02)00746-8

Walker, M.P., Brakefield, T., Hobson, J.A. & Stickgold, R. (2003). Dissociable stages of human memory consolidation and reconsolidation. *Nature*, 425, 616-620. doi:10.1038/nature01930

Walker, M.P. Brakefield, T., Seidman, J., Morgan, A., Hobson, J.A. & Stickgold, R. (2003). Sleep and the Time Course of Motors Skill Learning. *Learning & Memory*, 10, 275-284. doi: 10.1101/lm.58503

Walker, M.P. & Stickgold, R. (2004). Sleep-Dependent Learning and Memory Consolidation. *Neuron*, 44 (1), 121-133. doi: 10.1016/j.neuron.2004.08.031

Walker, M.P. (2005). A refined model of sleep and the time course of memory formation. *Behavioral and Brain Sciences*, 28, 51-64. doi: 10.1017/S0140525X05000026

Walker, M.P. (2009). The Role of Sleep in Cognition and Emotion. *The Year in Cognitive Neuroscience*, 168 – 197. doi: 10.1111/j.1749-6632.2009.04416.x

Wilhelm, I., Rose, M., Imhof, K.I., Rasch, B., Büchel, C. & Born, J. (2013). The sleeping child outplays the adult's capacity to convert implicit into explicit knowledge. *Nature Neuroscience*, 16 (4), 391 – 393. doi: 10.1038/nn.3343

Wrisberg, C.A. & Anshel, M.H. (1993). A field test of the activity-set hypothesis for warm – up decrement in an open skill. *Research Quarterly for Exercise and Sport*, 64 (1), 39-45. doi: 10.1080/02701367.1993.10608777

Wulf, G. & Shea. C.H. (2002). Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychonomic Bulletin Review*, 9 (2), 185-211. doi: 10.3758/BF03196276

Yokoi, A., Bai, W. & Diedrichsen, J. (2017). Restricted transfer of learning between unimanual and bimanual finger sequences. *Journal of Neurophysiology*, 117, 1043-1051. doi: 10.1152/jn.00387.2016

Danksagung

An dieser Stelle möchte ich ganz besonders meinem Betreuer, Prof. Dr. Klaus Blischke, danken. Und das insbesondere für die Unterstützung über den kompletten Zeitraum dieser Arbeit sowie die stete Ermutigung, eigene Ideen zu verfolgen. Unsere Treffen und Gespräche haben mir häufig neue Perspektiven eröffnet und mich dazu inspiriert, die nächsten Schritte in Angriff zu nehmen.

Verzeichnis dissertationsrelevanter Publikationen

Publikation 1:

Malangré, A., Leinen, P. & Blischke, K. (2014). Sleep-related offline learning in a complex arm movement sequence. *Journal of Human Kinetics*, 40 (1), 7 – 20. doi: 10-2478/hukin-2014-0002

Publikation 2:

Blischke, K. & Malangré, A. (2016). Chunk Concatenation Evolves with Practice and Sleep-Related Enhancement Consolidation in a Complex Arm Movement Sequence. *Journal of Human Kinetics*, 51 (1), 5- 17. doi: 10.1515/hukin-2015-0163.

Publikation 3:

Malangré, A. & Blischke, K. (2016). Sleep-Related Offline Improvements in Gross Motor Task Performance Occur Under Free Recall Requirements. *Frontiers in Human Neuroscience*, 10. doi: 10.3389/fnhum.2016.00134

Publikation 4:

Blischke, K. & Malangré, A. (2017). Task Complexity Modulates Sleep-Related Offline Learning in Sequential Motor Skills. *Frontiers in Human Neuroscience*, 11:374. doi: 10.3389/fnhum.2017.00374

Weitere Publikationen

Blischke, K., Erlacher, D., Kresin, H., Brückner, S. & Malangré, A. (2008). Benefits of Sleep in Motor Learning - Prospects and Limitations. *Journal of Human Kinetics*, 20 (1), 23 – 35. doi: 10.2478/v10078-008-0015-9

Vorträge und Konferenzbeiträge mit Review des Abstracts

Blischke, K. & Malangré, A.⁹ (2015). Sleep enhances chunk concatenation in a complex arm movement sequence. Vortrag auf der 14. Zwei-Jahrestagung der Sektion „Sportmotorik“ der Deutschen Vereinigung für Sportwissenschaft (dvs), „*The Athlete's Brain: Neuronale Aspekte motorischer Kontrolle*“, vom 05. bis 07. März 2015 in München, Deutschland.

Blischke, K. & Malangré, A. (2014). Sleep and gross-motor sequence production. Vortrag auf der 19. Jahrestagung des European College of Sport Science (ECSS) vom 02. bis 05. Juli 2014 in Amsterdam, Niederlande.

Malangré, A. & Blischke, K. (2014). Sleep boosts chunk concatenation in a gross motor task. Vortrag auf der 19. Jahrestagung des European College of Sport Science (ECSS) vom 02. bis 05. Juli 2014 in Amsterdam, Niederlande.

Blischke, K., Malangré, A. & Leinen, P. (2013). Sleep-related offline-learning and task complexity in a motor skill requiring redundancy control. Posterbeitrag auf der Annual Conference of the North American Society for the Psychology of Sport and Physical Activity (NASPSA) vom 13. bis 15. Juni 2013 in New Orleans (LA), USA.

Blischke, K., Malangré, A., Leinen, P. & Brueckner, S. (2011). Consolidation-based enhancement (CBE) in motor skill learning: State- or time-dependent? Vortrag auf der Annual Conference of the North American Society for the Psychology of Sport and Physical Activity (NASPSA) vom 11. bis 13. Juni 2011 in Burlington (VT), USA.

Malangré, A., Leinen, P. & Blischke, K. (2011). Consolidation based enhancement: State- or time-dependent? Vortrag auf der 12. Zwei-Jahrestagung der Sektion „Sportmotorik“ der Deutschen Vereinigung für Sportwissenschaft (dvs), „Embodiment: Wahrnehmung - Kognition – Handlung“, vom 20. bis 22. Januar 2011 in Köln, Deutschland.

Blischke, K., Erlacher, D., Kresin, H., Brückner, S. & Malangre, A. (2009). Differential effects of sleep on motor learning: A multitask approach. Vortrag auf der Annual Conference of the North American Society for the Psychology of Sport and Physical Activity (NASPSA) vom 11. bis 13. Juni 2009 in Austin (TX), USA.

⁹ Anteile an den jeweiligen Vorträgen: Bei Erstautorenschaft: Präsentation des Vortrages auf der jeweiligen Konferenz. Bei Mitautorenschaft (Nennung an zweiter, dritter oder vierter Stelle): Beitrag zur Vorbereitung der Präsentation.



Sleep-Related Offline Learning in a Complex Arm Movement Sequence

by

Andreas Malangré¹, Peter Leinen¹, Klaus Blischke¹

Sleep is known to elicit off-line improvements of newly learned procedural skills, a phenomenon attributed to enhancement consolidation of an internal skill representation. In the motor domain, enhancement consolidation has been reported almost exclusively for sequential-finger-tapping skills. The aim of the present study was to extend the notion of sleep-related enhancement consolidation to tasks closer to everyday motor skills. This was achieved by employing a sequence of unrestrained reaching-movements with the non-dominant arm. Fifteen reaching-movements had to be executed as fast as possible, following a spatial pattern in the horizontal plane. Terminating each movement, a peg had to be fitted into a hole on an electronic pegboard. Two experimental groups received initial training, one in the evening, the other one in the morning. Subsequently, performance in both groups was retested twelve, and again 24 hrs later. Thus, during retention each individual experienced a night of sleep, either followed or preceded by a wake interval. Performance error remained low throughout training and retests. Yet mean total execution time, indicative of task execution-speed, significantly decreased for all individuals throughout initial training (no group differences), and significantly decreased again in either group following nocturnal sleep, but not following wake. This finding does not appear to result merely from additional practice afforded at the time of retests, because only after a night of sleep individuals of both experimental groups also revealed performance improvement beyond that estimated from their initial training performance.

Key words: motor learning, sleep, memory consolidation, motor sequence, gross motor skill.

Introduction

Today, there is a considerable body of research in the neurobehavioral sciences addressing the consolidation and optimization of internal representations in the course of motor learning (Song, 2009). In particular it has been shown that after initial practice of a motor skill (and in the absence of any further physical practice) the elapse of time stabilizes performance, while sleep or daytime naps administered during the retention interval result in an additional performance enhancing effect. During the last decade this phenomenon, usually referred to as “offline learning” or “enhancement consolidation” (EC), has been corroborated by

numerous behavioral studies (Fischer et al., 2005; Walker, 2005).

However, nearly all of these studies employed very similar types of tasks frequently used in the domain of motor sequence learning: namely the serial-reaction-time task (SRTT) or the sequential-finger-tapping task (Albouy et al., 2013; Friedman and Korman, 2012). And although it was proclaimed ten years ago already that “this finding of sleep-dependent motor skill improvement may have important implications for the efficient learning of *all* skilled actions in humans” (Walker et al., 2002), even today there is still little evidence to support that claim. This is the more surprising, since practitioners in the

¹ - Institute of Sport Science, Saarland University, Saarbruecken, Germany.

applied field of movement studies (i.e. coaches and therapists) would readily introduce the notion of sleep-related offline learning to their practice routines, was it only confirmed with respect to tasks immediately relevant to their vocational domains. So the question still needs to be answered, as to what extent the above findings also apply to motor tasks beyond sequential-finger-tapping skills.

In the first attempt to systematically address this issue, we conducted a series of experiments involving different motor criterion tasks (Blischke et al., 2008). In these experiments, sleep-related offline learning in the standard sequential-finger-tapping task was successfully replicated. By contrast, sleep-related EC was neither observed (a) in a relative timing task incorporating an integer rhythm, nor (b) in a pursuit-tracking task with subjects being *unaware* of the spatial pattern to be learned, nor (c) when subjects had to precisely (re)produce a sub-maximal force impulse in a counter movement jump. While EC in rhythmic movements still is a matter of debate (Lewis et al., 2011; Wright et al., 2012), it has been argued elsewhere that sleep-related EC essentially depends on explicit (sequence) knowledge and awareness (Robertson et al., 2004). This plus assuming force impulses being represented in a rather implicit fashion, may well account for the absence of offline learning in tasks like pursuit-tracking and ballistic force production.

So at first perusal it seems that sleep-related EC requires some involvement of *declarative* memory processes, often associated with routines of *explicit* learning. Moreover, as perceptual skill learning (like in visuo-motor adaptation tasks) has been found to be rather sleep-independent (Doyon et al., 2009; but see Huber et al., 2004), sleep-related EC should be most pronounced in movement *sequences*, organized in *allocentric space* (Albouy et al., 2013; Cohen et al., 2005; Witt et al., 2010). Namely, fast and precise sequence execution requires rapid *in-advance specification* of requisite elements to be organized in immediate succession (Rhodes et al., 2004; Verwey, 1996). This process would be facilitated if spatial memory reorganization occurred across a period of sleep (Kuriyama et al., 2004). If all this holds true, then also *gross* motor tasks relevant to sports, occupational therapy, and motor rehabilitation should be amenable to sleep-related EC, if they

were only *spatially* defined, *sequentially* organized, and *explicitly* acquired.

This notion was initially tested in a study by Schmidt et al. (2010). Following a SRTT-paradigm, subjects were required to repeatedly produce a sequence of unrestrained arm movements: on a vertically positioned smart board, four horizontally aligned rectangles had to be touched with the hand(s) as fast as possible, thereby following a certain order. Actually, this task was an enlarged version of the same spatial pattern incorporated in finger-tapping tasks, which repeatedly had been proven subject to sleep-related EC when being executed on a key-board (Walker, 2005; Blischke et al., 2008). Two experiments were run on the smart board, with participants using either one or both hands. However, in contrast to the key-pressing studies, *no* sleep-related offline learning was found in any of these experiments incorporating gross limb movements instead of just activating different digits in a certain order (Schmidt et al., 2010).

At this point one might conclude that findings on sleep-related EC in sequential finger-tapping skills in fact do not generalize to gross motor tasks at all, suggesting some principal dissociation of fine and gross motor sequence representation. However, considering the impact motor skill complexity may have on sleep-dependent learning (Kuriyama et al., 2004), perhaps the arm-movement sequence in Schmidt and colleagues' study was just lacking the necessary amount of task difficulty: first, task complexity was low in general as the sequence incorporated just five elements. Also, precision requirements were insignificant considering the generous dimensions of the spatial goals on the smart board. With cognitive requirements and motor control demands being that low, any actual memory enhancement therefore might not have come overtly into effect at sequence recall tests. Moreover, as was suggested by individual reports, performance improvements due to EC could have been masked by muscular fatigue as subjects had to continuously produce rapid full-range arm movements in the frontal plane for 30 seconds at a time. Thus with increasing skill expertise and execution speed, subjects' performance might have approached a physically determined ceiling effect.

As a consequence, another attempt at

investigating sleep-related EC in gross-motor tasks was undertaken by introducing modified task conditions. In the study presented here, an unrestrained arm-movement sequence to be executed as fast as possible was employed again. However, to increase task difficulty the sequence this time comprised a series of 15 reaching movements (i.e. sequence elements) with the (non-dominant) hand serving as a single end-effector. Also, drawing on the systematic effect of target width and movement amplitude on movement time first mathematically analyzed by Fitts (1954), precision requirements for all sequence elements were set at an index of difficulty (ID) of 4.95 on average, with IDs > 4.5 being regarded as high (Boyle and Shea, 2011). According to Fitts, the ID is determined by the equation $\text{Log}_2(2A/W)$, where A represents the movement amplitude measured from one target center to the other target center and W represents the width of the target area in the direction of the movement. Furthermore, in order to somewhat reduce muscular fatigue the movements this time had to be carried out in the horizontal plane. Also, unlike in the SRTT-paradigm each sequence execution trial now was triggered by a separate start signal. So the present task to some extent resembled features of a discrete sequence production task (DSPT; Rhodes et al., 2004), thereby reducing once more the chance of fatigue building up within blocks of successive trials. Altogether the gross-motor task employed this time appears to be largely cleared of those features which possibly diminished effects of sleep-related EC in earlier studies, yet at the same time still bears good resemblance to sport skills and activities of daily living.

Thus, with respect to this gross-motor task comprised by a sequence of unrestrained arm-movements, it was hypothesized that after initial learning *sleep*, but not *wake*, significantly facilitates performance (namely: execution speed) at retention beyond any improvement merely being afforded by exposure to additional practice trials.

Material and Methods

Subjects

24 subjects (25.5 ± 3.9 years; 6 females, 1 left-handed; 18 males) participated in this study, which was conducted in accordance with the

ethical standards of the 1964 Declaration of Helsinki. All participants gave their written informed consent. Subjects were randomly assigned to two experimental groups, the EME (= Evening-Morning-Evening)-group (N = 12) and the MEM (= Morning-Evening-Morning)-group (N = 12). Sample sizes were chosen so that possible performance enhancements at retention should turn out significant if they were at least the size of those encountered previously for the standard finger-tapping task, which yielded η^2_p -values of $\geq .326$ in the MEM-groups and $\geq .519$ in the EME-groups for improvements in execution speed (Blischke et al., 2008). Since previous studies showed that performance is unaffected by gender, experimental groups were not balanced with respect to sex of participants. Participants were required to refrain from daytime naps, alcohol, excessive caffeine-intake, and any other drugs from the night before training session until the end of the experiment. Physical activity (e.g. sport practice) was permitted.

Task and dependent measures

Holding a small peg with their non-dominant hand, on a start signal subjects were to carry out a sequence of 15 arm movements in the horizontal plane. Following a fixed pattern of end-point locations, these movements differed in range (from 3.83 to 33.75 cm) as well as in direction (Picture 1, lower panel). With their hand visible all the time, at the end of each movement participants had to quickly fit the peg into a hole of a pegboard in front of them, thereby closing a magnetic contact. The pegboard employed consisted of two horizontal wooden bars (41.7 cm long, 16 cm apart), each containing ten holes 22.22 mm in depth, 12.7 mm in diameter, and 25.4 mm apart in the left-right dimension, 195 mm apart in the forward-backward dimension (Picture 1). On a computer screen representing the 20 pegboard holes, each present movement goal was illuminated red until the respective magnetic contact was closed, turning the color from red to green. At the same time the next goal was indicated by turning to red. Once the respective sequence element was terminated, the next movement had to be started immediately, until the sequence was completed. No additional information (e.g. augmented feedback) was provided. As participants could use their whole arm by freely moving shoulder, elbow and wrist,

they were required to control redundant biomechanical degrees of freedom. Subjects were instructed to carry out each single sequence-trial as fast and with as few errors as possible, and not to speed up performance at the risk of increasing number of errors. Dependent measures taken for each subject were *Total Execution Time* (TET) per sequence, averaged over the number of correct sequences per trial block, and the number of *Erroneous Sequences* (ES) per trial block. Thus, TET is inversely proportional to sequence execution speed.

Design and procedure

After being shortly familiarized with the electronic pegboard and the peg-plugging procedure in general, both experimental groups received initial training of the criterion task (ten blocks of ten trials each), the EME-group in the evening (7 to 9 p.m.), the MEM-group in the morning (7 to 9 a.m.). Both groups then were retested 12 hrs (Retest 1), and once again 24 hrs later (Retest 2), with each Retest comprising three blocks of ten trials. Thus, subjects in the EME-group had a regular night's sleep during their first, those in the MEM-group during their second retention interval. Trial blocks always were separated by a pause of 30 seconds during practice as well as in Retests.

Statistics

First, for each subject ES- and TET-measures were averaged across trials per block. Then for each subject and dependent variable, a "Post-Training" measure was calculated from the last three initial training blocks (blocks 8, 9 & 10), while Retest 1- and Retest 2-measures were calculated from blocks 11, 12 & 13, and 14, 15 & 16, respectively. Group means were calculated on this basis. For inferential statistics, two-way ANOVAs on the factors "Group" x "Acquisition-Block" and "Group" x "Test" (levels: Post-Training, Retest 1, Retest 2), one-way ANOVAs on the factor "Test", and paired *t*-tests were run. With respect to repeated-measures factors, in case of violation of the sphericity assumption *df*-correction according to Greenhouse-Geisser was applied. A significance level of $p \leq .05$ was used for all inferential statistics. Calculations were conducted with SPSS-PC, version 15.0. Effect sizes were provided in terms of η^2_p with respect to ANOVAs, and Cohen's *d* with respect to *t*-tests.

Results

Performance during acquisition

To assess changes in performance during initial training, 2[Group] x 10[Block]-ANOVAs (repeated-measures factor "Block") were calculated on the respective data (Figures 1 and 2). On the average, there are 2.8 (± 2.01) *Erroneous Sequences* (ES) per trial block (EME: 3.08 ± 2.1 ; MEM: 2.56 ± 2.16). Error rate is small from the beginning and does not fluctuate significantly across acquisition ($p = .079$, $\eta^2_p = .085$). There is no Group x Block interaction ($p = .273$, $\eta^2_p = .054$), nor do groups differ in performance ($p = .374$, $\eta^2_p = .036$). In contrast, *Total Execution Time* (TET) significantly decreases during acquisition in both experimental groups ($F_{[Block]}$ (4.304, 94.690) = 98.053, $p < .001$, $\eta^2_p = .817$). Again there is no Group x Block interaction ($p = .295$, $\eta^2_p = .054$), nor do groups differ ($p = .832$, $\eta^2_p = .002$).

Comparing performance at the end of practice and retention

To assess possible changes in performance across the two retention intervals, 2[Group] x 3[Test] ANOVAs with "Test" as a repeated-measures factor (levels: Post-Training, Retest 1, Retest 2) were calculated on the respective data (Figures 1 and 2). In accord with our central hypothesis, significant Group x Test interactions (in combination with significant pairwise comparisons concerning the interaction term) were taken as evidence for sleep-related EC in a particular retention interval if, according to the descriptive data for those participants who had slept during that interval, improvements in performance were indeed present and at the same time larger than those for subjects who had stayed awake during the same period of time.

In such a case, however, in order to unambiguously attribute possible changes in performance to the exact succession of wake and sleep periods, data were also analyzed separately for each circadian condition (EME and MEM) by means of repeated-measures one-way-ANOVAs on the factor "Test". Only this way it is possible to unequivocally dissociate within each group periods reflecting true performance improvements from periods with performance remaining more or less the same, thus making allowance for the fact that by theory (Walker, 2005) enhancement consolidation is attributed to sleep, but not just to the elapse of time.

Performance errors

According to a 2[Group] \times 3[Test] ANOVA, the overall number of Erroneous Sequences (ES) does not change across tests at all ($p_{[\text{Test}]} = .239$, $\eta^2_p = .063$), nor do groups differ on the whole ($p_{[\text{Group}]} = .904$, $\eta^2_p = .001$). There is, however, a significant Group \times Test interaction ($F(1.575, 34.658) = 4.996$, $p = .018$, $\eta^2_p = .185$). Also, each of the respective pairwise comparisons turns out to be significant (Post-Training, Retest 1: $F(1, 22) = 7.650$, $p = .011$, $\eta^2_p = .258$; Retest 1, Retest 2: $F(1, 22) = 5.067$, $p = .035$, $\eta^2_p = .187$). However, considering the somewhat ambiguous descriptive data (Figure 1), even in the light of these results the following questions still remained to be answered: (a) is error reduction observed in the EME-group following sleep at Retest 1, which is indicative of sleep-related EC, completely lost again during following wake-interval? And, (b) does error rate in the MEM-group change at all? Thus for further clarification, repeated-measures one-way-ANOVAs on the factor "Test" were run for each of the experimental groups separately.

As is corroborated this way, the error rate significantly changes during retention in the EME-group only ($F_{[\text{Test}]}(2, 22) = 4.948$, $p = .017$, $\eta^2_p = .310$). According to the respective pairwise comparisons (i.e. within-subjects contrasts), for this group a significant reduction in errors occurs during the *first* 12-hr retention interval, i.e. following *sleep* (Post-Training, Retest 1: $F(1, 11) = 12.008$, $p = .005$, $\eta^2_p = .522$). During the following wake period the error rate increases somewhat again indeed, yet without this effect reaching statistical significance (Retest 1, Retest 2: $p = .088$, $\eta^2_p = .242$). However, in the MEM-group, the respective ANOVA covering the total 24-hr retention interval does not yield significance altogether ($p = .332$, $\eta^2_p = .091$).

Execution time

According to a 2[Group] \times 3[Test] ANOVA, Total Execution Time (TET) significantly decreases across tests in both groups ($F_{[\text{Test}]}(2, 44) = 45.590$, $p < .001$, $\eta^2_p = .675$), while groups as such do not differ ($p_{[\text{Group}]} = .843$, $\eta^2_p = .002$). Again there is a significant Group \times Test interaction ($F(2, 44) = 4.308$, $p = .020$, $\eta^2_p = .164$). According to the respective pairwise comparisons, this interaction pertains as well to the first ($F(1, 22) = 6.071$, $p = .022$, $\eta^2_p = .216$) as to the second 12-hr retention interval ($F(1, 22) = 11.442$, $p = .003$, $\eta^2_p = .342$).

Considering the descriptive data (Figure 2), these results strongly support our central assumption of sleep-related EC coming into place in both experimental groups in terms of significantly shorter TET, that is increased execution speed. At this point it remains unclear, however, if TET in both groups also decreases significantly, though to a lesser extent, during their respective wake intervals. Thus, again, repeated-measures one-way-ANOVAs on the factor "Test" were run for each of the experimental groups separately.

Results are as follows: throughout the total 24-hrs retention period, TET significantly decreases in the EME-group ($F_{[\text{Test}]}(1.315, 14.468) = 19.904$, $p < .001$, $\eta^2_p = .644$). As pairwise comparisons indicate, this is essentially due to considerable off-line improvements following the *first* 12-hr retention interval, i.e. after this group's *sleep* period (Post-Training, Retest 1: $F(1, 11) = 35.112$, $p < .001$, $\eta^2_p = .761$), while during the subsequent wake interval TET in this group does not indicate any significant alteration (Retest 1, Retest 2: $p = .150$, $\eta^2_p = .179$). In the MEM-group, TET also significantly decreases throughout the total 24-hrs retention period ($F_{[\text{Test}]}(2, 22) = 32.910$, $p < .001$, $\eta^2_p = .749$). And again, according to the pairwise comparisons, TET does not change significantly throughout this group's wake interval (Post-Training, Retest 1: $p = .127$, $\eta^2_p = .199$), but significantly decreases throughout its *sleep* period, which is the MEM-group's second retention interval (Retest 1, Retest 2: $F(1, 11) = 60.337$, $p < .001$, $\eta^2_p = .846$) (Figure 2 and 3).

Comparing actual and estimated Retest-measures (execution time)

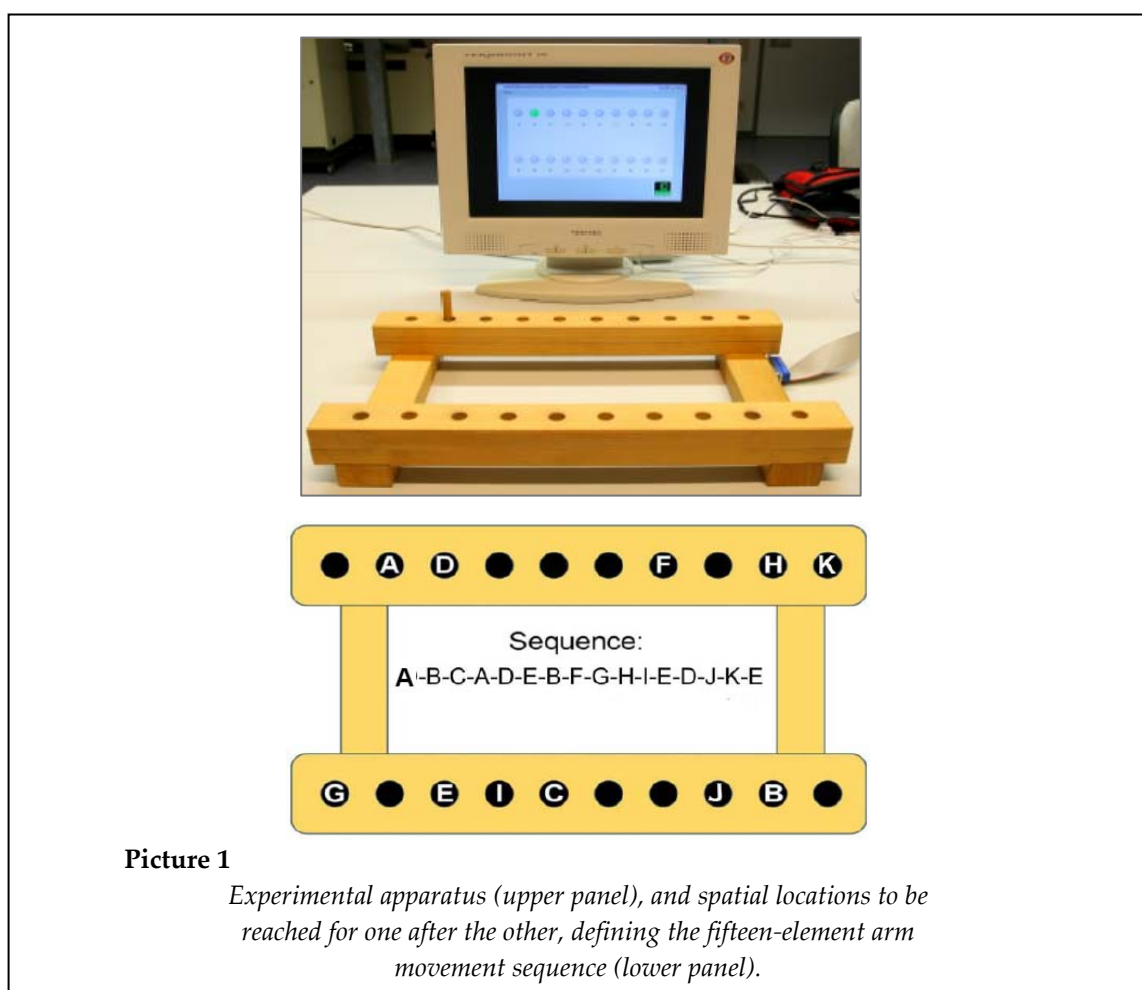
As was shown above, TET shows considerable and continuous improvement throughout acquisition in either group. Consequently, and different from the error rate, it has to be considered that TET may not have fully reached asymptotic performance at the end of initial training. In fact, in the EME-group TET decreases until the very last block of initial practice (Figure 2, left panel), while in the MEM-group, TET leveling off from block 9 to 10 might just reflect one of several transient performance fluctuations typical for this group (Figure 2, right panel: blocks 2 & 3, blocks 4 & 5). Therefore, continued practice, uninterrupted by retention intervals, possibly might have caused further improvements in performance quite similar to those actually

reported at recall following sleep. Thus, it had to be determined, whether the significant reduction in TET (i.e. improvement in execution speed) for individuals of both experimental groups was indeed a result of sleep-related EC, or just a consequence of further practice.

One way to assess EC considering continued learning at retention, involves extrapolation of each subject's respective initial training data. These *estimated* retention data are then used in conjunction with the individuals' *actual* performance on the Retest trials. If the actual performance is better (i.e. TET lower) than the predicted performance, offline facilitation is assumed to have occurred. In the present study, estimated TET-Retest-measures were provided as follows: based on each single subject's TET-acquisition data (means per trial block), for each individual a power function of the type $y = kn^c$ was calculated and used to obtain an estimate for that individual's performance for the additional six trial blocks during both Retest 1 and Retest 2. Power functions are widely used to

mathematically model practice-dependent changes in performance in the course of skill acquisition (Newell and Rosenbloom, 1981; Newell et al., 2009). Estimated TET-data for each individual then were collapsed across blocks in each test, thus providing mean predicted TET-performance at Retest 1 and Retest 2 for each one subject.

Thus, if sleep (but not wake) had indeed enhanced memory consolidation, *actual* TET should turn out significantly *lower* as compared to estimated TET in either group when tested right after the respective retention interval filled by sleep (Retest 1, EME-group; Retest 2, MEM-group). Also, we expected sleep-related performance facilitation to be preserved throughout an additional wake-period, which can be tested for at Retest 2 in the EME-group. According to these *a-priori* hypotheses, for each group one-tailed paired *t*-tests were calculated in order to compare actual and estimated TET-data at Retests.



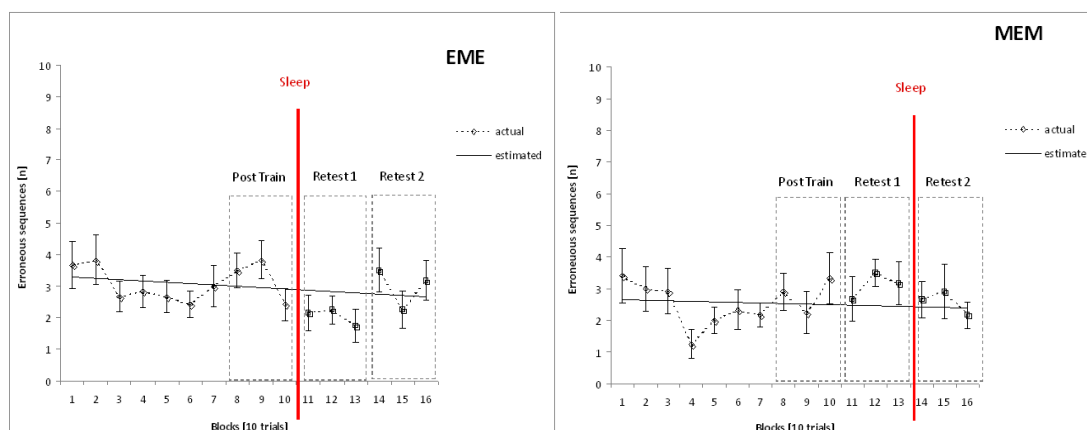


Figure 1

Number of Erroneous Sequences (ES) per trial block: initial training (blocks 1 through 10) and subsequent retests. Symbols represent group means per trial block (actual performance); error bars: standard errors of the mean.

Solid lines represent linear functions derived from group mean initial training data (EME: $y = -0.042x + 3.31$; MEM: $y = -0.018x + 2.66$).

Left panel: EME-group; right panel: MEM-group.

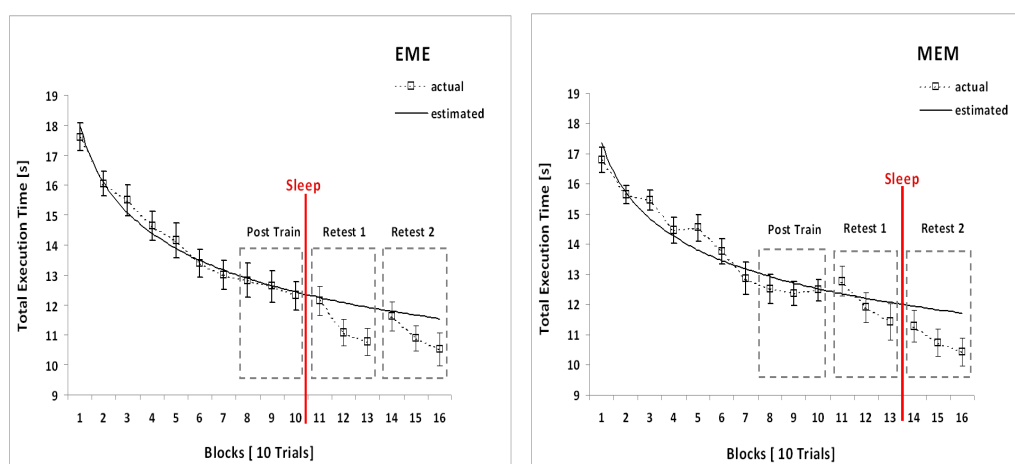


Figure 2

Mean Total Execution Time (TET; seconds) per trial block (correct sequences only): initial training (blocks 1 through 10) and subsequent retests.

Symbols represent group means per trial block (actual performance);

error bars: standard errors of the mean. Solid lines represent power

functions derived from the group mean initial training data (EME: $y = 17.97x^{-0.16}$, $R^2 = 0.98$;

MEM: $y = 17.36x^{-0.142}$, $R^2 = 0.92$). Left panel: EME-group; right panel: MEM-group.

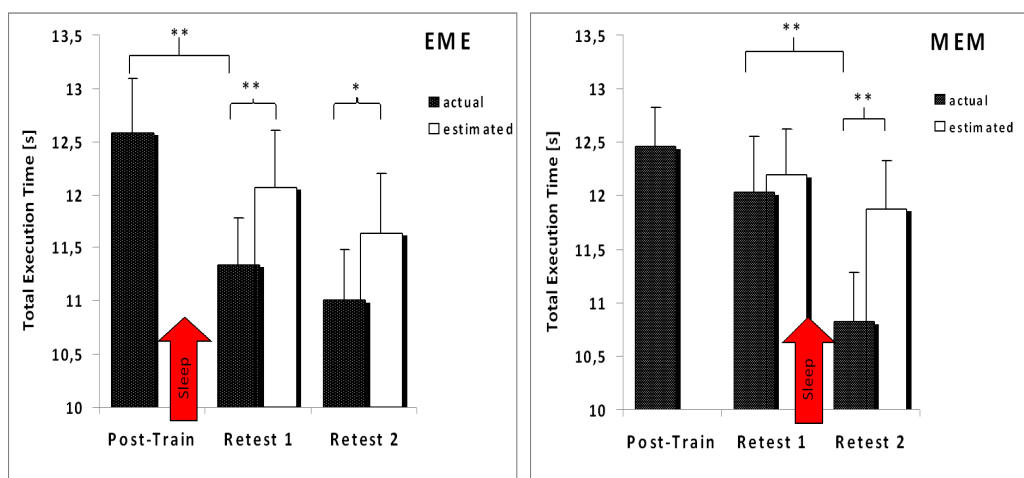


Figure 3

Total execution time (TET) at Post-Training (including block 8, 9 & 10), at Retest 1 (including block 11, 12 & 13) and at Retest 2 (including block 14, 15 & 16). Filled bars: actual data; open bars: estimated data. Presented are group means per test. Error bars: standard error of the mean. Left panel: EME-group; right panel: MEM-group.

As can be inferred from Figure 3, in each group the observed *actual* TET proves to be *significantly shorter* than the predicted TET *only after sleep* ($\text{EME}_{[\text{Retest 1}]}: t(11) = -3.901, p_{[\text{one-tailed}]} = .001, d = 1.13$; $\text{MEM}_{[\text{Retest 2}]}: t(11) = -5.019, p_{[\text{one-tailed}]} < .001, d = 1.41$). Also, in the EME-group this sleep-induced advantage of actual over estimated TET-performance appears to be preserved during the wake interval following sleep ($\text{EME}_{[\text{Retest 2}]}: t(11) = -1.967, p_{[\text{one-tailed}]} = .038; d = .57$). In the MEM-group, however, actual and estimated TET-performance are not dissociated at all by the wake interval *preceding* sleep ($\text{MEM}_{[\text{Retest 1}]}: t(11) = -.825, p_{[\text{two-tailed}]} = .427, d = .25$). According to these results, the significant reductions in TET reported for both experimental groups in the course of a 24-hr retention period cannot be attributed to merely continuing practice. Rather, and in support of our central hypothesis, they provide evidence for true sleep-related *offline-learning*, independently of sleep being administered during the first or during the second half of a 24-hrs retention interval.

Discussion

As outlined in the introduction, previous findings on sleep-related EC in motor skill

learning appear to be restricted almost exclusively to a specific set of fine motor tasks, namely sequential-finger-tapping skills (Albouy et al., 2013; Cohen et al., 2005; Doyon et al., 2009; Friedman and Korman, 2012; Fischer et al., 2005; Kuriyama et al., 2004; Lewis et al., 2011; Robertson et al., 2004; Walker et al., 2002; 2005; Witt et al., 2010). Earlier attempts at demonstrating sleep-related offline learning for motor tasks different from finger-tapping, for the most part came to nothing (Blischke et al., 2008). It was concluded, however, that at least certain *gross* motor tasks might be amenable to sleep-related EC, as long as they were sequentially structured, spatially defined, and explicitly acquired.

To test this notion, in the present experiment subjects practiced a complex sequence of 15 arm-movements with high spatial precision requirements. This task was executed on an electronic pegboard and involved a single end-effector (i.e. the non-dominant hand holding a small peg in a pincher-grip), while all three joints of the arm (i.e. wrist, elbow, and shoulder) were to move freely. End-point locations of the sequence's elemental reaching-movements formed a two-dimensional spatial pattern. Subjects were to avoid any sequence errors (ES),

and at the same time to minimize total execution time (TET) per trial. Visual guidance was provided during sequence execution on a computer screen, but there was no additional feedback on performance measures. Initial training of this task was administered either in the evening (EME-group) or in the morning (MEM-group). Subsequently, subjects in either group underwent two retests 12 (Retest 1) and then again 24 hrs (Retest 2) later. During retention, individuals in each group slept at night, and kept awake during the respective daytime interval. It was expected that sleep, but not wake, would enhance performance at retention beyond any improvement merely being afforded by exposure to additional practice trials.

As it turned out, *performance error (ES)* was rather low for individuals for both groups right from the start, and did not improve any more throughout initial training. Mean ES then significantly decreased during the sleep-filled retention interval (i.e. from Post-Training to Retest 1) in the EME-group only, but remained unchanged throughout the total 24-hr retention period in the MEM-group. So there is only partial evidence for sleep-related EC on account of the error data, but this latter finding may not be surprising given the small number of errors over all. Obviously, subjects closely followed instructions to keep the error rate at bay, and not to increase execution speed at the cost of increasing number of errors at the same time. In fact have error measures (as opposed to the respective speed measures) repeatedly been proven to be less or even non-sensitive to sleep-related EC also in sequential-finger-tapping tasks (Albouy et al., 2013; Doyon et al., 2009). In any case the present results clearly indicate that the between-session changes in performance speed referred to below did not occur at the expense of performance accuracy.

In contrast *Total Execution Time (TET; non-erroneous sequences only)*, an index of *performance speed*, showed considerable improvement for all individuals in each group throughout initial training without ever leveling off (no group differences), thus rendering continued learning even at retests highly possible. During retention, mean *actual TET* in each group *significantly decreased* again, when being retested *following sleep* (i.e. at Retest 1 in the EME-group; at

Retest 2 in the MEM-group), but always remained stable throughout the respective wake interval. Although this result appeared to be well in line with the notion of sleep-related EC, in order to arrive at a definite conclusion here continued learning at retests first had to be ruled out as an alternative interpretation of these data. To this end, TET measures to be expected at retests if practice had continued uninterrupted by retention intervals were estimated by means of power functions derived from each individual's initial training data. These predicted TET-data then were compared to the actual TET-data established at retests, which procedure yielded the following results: when retested *following sleep*, in both groups *actual TET* turned out *significantly shorter* than *predicted TET*, while both measures did not differ statistically just following wake in the MEM-group. This result confirms that actual TET-reductions found at retests following sleep, cannot be attributed merely to continued learning, but rather reflect some EC.

Given the fact that visual guidance was provided during sequence execution in our experiment, the present results might reflect not only effects of sequence memory enhancement, but also of sleep-related improvements of mechanisms relevant to the online processing of visual stimulus information. Our results could be more purely related to EC of the sequence representation if directly after the practice sessions as well as during retests following the respective retention intervals, subjects had also produced the arm movement sequence without any visual instruction stimuli in a free recall condition, and yielded the same results. We did not introduce such a testing procedure in our present study in order to avoid possible confounds if the same subjects underwent different retest conditions (i.e. with and without visual guidance). However, we have recently applied this procedure to a follow-up study, which is still in progress. From this study, twelve subjects (all young adults) so far have been analyzed. They initially practiced a ten-element arm movement sequence on the peg-board in the morning under visual guidance (120 trials), and then reproduced the same sequence under free recall conditions 15 minutes (Post-Training), 12 hrs (Retest 1) and – after a night of sleep – 24 hrs later (Retest 2). Eight participants were able to

freely recall the sequence even 24 hrs after acquisition. Preliminary results on these eight subjects show about the same pattern of sleep-related EC with respect to TET as reported in the present paper (Post-Training, Retest 1: $p = .922$, $\eta^2_p = .001$; Retest 1, Retest 2: $F(1, 7) = 8.269$, $p = .024$, $\eta^2_p = .542$). Also, performance in the free recall early retention test (i.e. Post-Training) was not different from that during the last three blocks of practice with visual sequence information provided ($p = .919$, $\eta^2_p = .002$). We therefore conjecture the findings of EC presented in the present paper also to express over-night enhancement of *sequence memory* rather than facilitation of online processing of visual stimulus information.

It has been questioned previously whether performance improvements following sleep are real or whether such changes rather reflect time-of-day confounds (e.g. compensation of fatigue effects) due to the experimental protocols (Brawn et al., 2010; Cai and Rickard, 2009). However, in the present study the extent of learning during acquisition and performance levels at the end of training were quite similar in both groups, regardless of whether the training took place in the evening or in the morning. This then suggests that the expression of delayed sleep-related post-training gains reported here are indeed due to the EC process rather than the mere result of fatigue compensation taking place over night. This does not completely rule out circadian confound, however, as one might argue the EC process to be related rather to circadian influences than to the experience of sleep as such. All the same, empirical evidence to date does not support any such notion: first of all, there is abundant evidence for EC being correlated either with certain sleep parameters or (in general to a lesser extent) with the elapse of time (Fischer et al., 2005; Walker, 2005), while there is no indication whatsoever of certain circadian (i.e. time-of-day) aspects being of relevance here. Secondly, sleep-related EC has repeatedly been shown for the serial-finger-tapping task following day-time sleep, thereby incorporating experimental paradigms better suited for controlling for the time-of-day difference between acquisition and test sessions (Albouy et al., 2013; Doyon et al., 2009). And, last but not least, we found an increase in performance in both experimental groups only

after sleep but not after the wake period, thus ruling out that EC was significantly influenced by retention interval duration.

One might also ask if the above findings on sleep-related EC might be biased by our participants' dexterity, since they carried out the present experiment with their non-dominant arm. In this respect in the present study we merely followed the standard procedure applied by the vast majority of studies on sleep-related EC. Surprisingly, in the literature surveyed we did hardly find any comment on this procedure, let alone an explicit justification. Yet, there is indeed a theoretical possibility that sleep may interact differentially with the hand or arm used during motor sequence learning. To date, however, there is hardly any evidence to support such an assumption. To our knowledge only one study so far involved two groups of subjects who practiced a novel finger opposition sequence either with their dominant or with their non-dominant hand, and were retested 24 hrs later (Balas et al., 2007). In general, results of that study suggested that training of either hand may trigger delayed gains in performance. Baseline data presented by Balas et al. (2007) for each of the two groups enrolled in their Experiment 1, however, do not rule out the possibility of slight advantages of the dominant over the non-dominant hand in this respect. Unfortunately this aspect was not subjected to any statistical proof in that study.

All in all, we believe our results to be well in accord with the above hypothesis, namely that even in a *gross* motor task involving a sequence of coordinated limb movements, sleep following initial learning significantly facilitates performance. This finding at the same time successfully extends the notion of sleep-related EC beyond the standard finger-tapping paradigm. It has to be stated, however, that the present finding only was achieved after deliberately increasing task difficulty (in terms of precision requirements) as well as task complexity (in terms of sequence length), as compared to an earlier attempt in scrutinizing sleep-related EC in an arm-movement sequence (i.e. Schmidt et al., 2010). However, at this point it cannot yet be determined to what extent either of these features (i.e. task complexity reflecting memory load, or spatial precision requirements reflecting motor control demands) specifically contributed to the

results presented here.

So undoubtedly, more and detailed information is needed on what task features precisely are essential for sleep-related EC to come into effect in gross motor skills. Only then could this concept truly be of relevance in the various applied fields of motor sciences. Some further leads that might help identifying task features relevant in this context could perhaps be derived from a very recent study by Kempler and Richmond (2012), which came to our knowledge only after the experiment presented here was completed already. Encouragingly these authors, too, found sleep-related offline learning in a gross motor task. Their task required participants to complete a sequence of six different body configurations, with each configuration defined by the two arms pointing simultaneously into the same or different directions (i.e. straight up or down or 90° sideways) in the frontal plane. So actually this task in a way mimicked a signalman transmitting a message through a series of flag signals. This sequence had to be repeated continuously over periods of 30 seconds during training and retests. Dependent measure was the average number of accurate (half)cycles (i.e. three correct configurations in a row) per each 30 s trial block as an index of speed.

Kempler and Richmond in their study did not report any quantitative criteria for correct angular positioning (i.e. up, down, or sideways) of their subjects' arms. Thus spatial precision requirements presumably were rather low, as was sequence length, which may have considerably reduced control demands as well as memory load of their task. And yet these authors found sleep-related offline learning all the same. This at first sight seems to be at odds with our above supposition, according to which EC can be reliably shown in gross motor tasks only if task difficulty is set at a sufficiently high level. However, while Kempler and Richmond followed the same experimental protocol as we did in our present experiment, sample sizes involved in their study were about three times those of ours (i.e. 35 as compared to 12 participants per experimental group). With sample sizes that large, even a seemingly small effect (due to only moderate task difficulty) might have turned out as statistically significant. Unfortunately, Kempler and Richmond did not report any effect sizes.

On the other hand, the Kempler and Richmond task required participants to *simultaneously* produce heterogeneous movements with both arms at transitions between the elementary sequence positions. This task feature perhaps induces a dual-task requirement as long as each arm's position has to be specified separately, a process that virtually promotes positioning errors or at least prolongs execution time, because at an early learning stage the positioning responses for both arms cannot yet be selected and prepared at the same time. This interpretation clearly draws on the concept of a "response selection/preparation"-bottleneck, a concept well established by researchers utilizing the *Psychological Refractory Period* (PRP) paradigm, which was originally introduced by Harold Pashler (Pashler and Johnston, 1998). At later learning stages and sufficient practice, however, the respective motor actions to be taken by either arm at each transition between sequence elements should eventually be integrated ("chunked") into one-and-the-same memory representation. Progressive chunk formation of this kind ought to reduce interference at recall, and help speeding up sequence execution (Rhodes et al., 2004; Verwey, 1996). To some extent, such internal chunking might also have taken place offline during the sleep-filled retention intervals, resulting in the respective sleep-related improvements of performance observed by Kempler and Richmond (2012).

Indeed, some evidence of sleep-related chunking has been provided by Kuryiama et al. (2004) already as far as the sequential-finger-tapping task is concerned. This was achieved by demonstrating significant overnight improvement of subjects' slowest key-press transitions, which was not observed during acquisition, nor following a day-time wake interval. Interestingly, these authors also found the greatest amount of sleep-related offline improvement (i.e. increased execution speed) for just that task configuration, which incorporated the highest level of task complexity (i.e. sequence length) and the greatest degree of between-limb coordination (i.e. bimanual task execution involving all eight digits). For an explanation, the authors argue that the cerebral network size of the sequence's memory representation, and thereby the potential for sleep-dependent synaptic plasticity, would

increase considerably with the number of digits (or in general end-effectors) involved in cooperative movement production.

As matters stand it cannot be decided whether Kempler and Richmond's findings are owing to their large sample sizes or rather to offline integration of elementary memory components and/or improvements in between-limb coordination. So with respect to gross motor tasks, the relation of certain task characteristics (like e.g. task difficulty, task complexity, and demands on between-limb coordination) and sleep-related EC is still up to discussion, as is the question to which extent chunking processes actually contribute to offline improvements in skilled motor performance (Kuriyama et al., 2004; Wright et al., 2010). While these questions are presently addressed by our group, certainly more research will be needed to this end.

Another point at issue relates to the *type* of motor sequence representation, which is supposed to be enhanced offline while subjects are asleep. According to a widely accepted model proposed by Hikosaka and his colleagues (1999), the acquisition of sequential behaviors resides in the interaction between different neural networks that would encode the same motor sequence in two different coordinate systems (i.e., spatial and motor). One memory component is thought to incorporate allocentric (spatial) coordinates, and to constitute an abstract effector-independent representation of a series of movements that need to be executed in an external frame of reference. The other memory component is supposed to be mediated through egocentric (motor) coordinates, and thus to constitute an effector-dependent, movement-based skill realized in an internal frame of reference (Hikosaka et al., 1999).

Hikosaka's model has also been applied to research on motor memory consolidation. For explicitly learned finger-tapping sequences it has been shown conclusively, that *sleep* specifically favors enhancement of the *extrinsic* (spatial) sequence representation, while consolidation of the respective *intrinsic* (motor) representation was *not* modulated by the sleep/wake condition (Albouy et al., 2013; Witt et al., 2010). While Hikosaka's integrative model proposes that the spatial component is created rapidly early during training with the motor component developing more slowly with extended practice, according to

Albouy and colleagues' results both spatial and motor representations exist already after minimal training, with only the former being amenable to sleep-related EC. However, the question arises as to what extent this dissociation also applies to gross motor tasks, where motor control requirements typically exceed those encountered in sequential-finger-tapping skills.

Furthermore, the laterality issue addressed above still remains of interest in the light of recent findings on the inter-limb transfer of multiple-element sequences. Here, non-dominant to dominant limb transfer has been shown to be superior to dominant to non-dominant limb transfer, if an additional load was added during acquisition, and if transfer tests required end-effector movement to the same spatial positions that had been practiced during acquisition (i.e. no mirror movements) (Panzer et al., 2010). For unloaded sequences, however, this type of effector transfer has been proven to be symmetrical (Kovacs et al., 2009). Asymmetric inter-limb effector transfer thus appears to be tied up with specific task requirements, here those concerning the control of movement dynamics. These and other findings (Sainburg, 2005) give rise to theoretical models of hemispheric specialization. In the end and with respect to specific task requirements, such models might also be of some relevance to concepts of motor memory consolidation.

It should be noted that polysomnographic data acquisition and analysis were not within the scope of the present study. Therefore, the present data do not allow for any conclusions regarding a particular sleep stage promoting EC of gross motor skill representation. With respect to the allocentric sequence representation of sequential-finger-tapping skills, there is increasing evidence for sleep-dependent gains in performance being correlated with the density of NREM sleep spindles and NREM stage 2 sleep duration (Albouy et al., 2013; Barakat et al., 2011; Walker et al., 2002; Witt et al., 2010). Other authors, however, have found this type of sequence representation (and declarative memory consolidation in general) to be correlated with REM sleep duration (Cohen et al., 2005; Fogel et al., 2007). Considering then gross motor tasks again, this question also calls for closer inspection in future research.

In conclusion, the present findings again successfully extend the notion of sleep-related EC to a gross motor task, i.e. a complex arm movement sequence. At the same time we are confident that our experimental paradigm is also well suited to address systematically and in some more detail some of those questions discussed here. While suchlike intents would be primarily related to the field of basic research, our findings also bear some practical implications relevant to sport and rehabilitation, whenever gross motor tasks have to be learned or relearned efficiently. This might specifically pertain to complex skills

involving the upper extremities and requiring fast and precise execution of a series of sub-movements in Euclidian space. Here, we would recommend practice to be distributed over two or more sessions separated by night or daytime sleep, respectively. Distributed practice schedules of this kind might be especially beneficial in situations, where the amount of practice per session is limited of necessity, as is often the case in rehab-training or acquisition of sport skills requiring extraordinary high physical or mental effort.

Acknowledgements

Part of this work was presented in a poster presentation at the annual conference of the North American Society for the Psychology of Sport and Physical Activity (NASPSPA), June 13th-15th, 2013, at New Orleans, LA, USA (Abstract published in conference proceedings: Blischke, K., Malangré, A., Leinen, P. (2013). Sleep-related offline-learning and task complexity in a motor skill requiring redundancy control. *Journal of Sport and Exercise Psychology*, 35, (Supplement), 21)

References

- Albouy G, Fogel S, Pottiez H, Nguyen VA, Ray L, Lungu O, Carrier J, Robertson E, Doyon J. Daytime sleep enhances consolidation of the spatial but not motoric representation of motor sequence memory. *PLoS ONE* 8, 2013; e52805. doi:10.1371
- Balas M, Netser S, Giladi N, Karni A. Interference to consolidation phase gains in learning a novel movement sequence by handwriting: dependence on laterality and the level of experience with the written sequence. *Exp Brain Res*, 2007; 180: 237-246
- Barakat M, Doyon J, Vandewalle G, Morin A, Poirier G, Martin N, Lafortune M, Karni A, Ungerleider LG, Benali H, Carrier J. Fast and slow spindle involvement in the consolidation of a new motor sequence. *Behav Brain Res*, 2011; 217: 117-121
- Blischke K, Erlacher D, Kresin H, Brueckner S, Malangré A. Benefits of sleep in motor learning – prospects and limitations. *Journal of Human Kinetics*, 2008; 20: 23-35
- Boyle JB, Shea CH. Wrist and arm movements of varying difficulties. *Acta Psychol*, 2011; 137: 382-396
- Brawn TP, Fenn KM, Nusbaum HC, Margoliash D. Consolidating the effects of waking and sleep on motor-sequence learning. *J Neurosci*, 2010; 30: 13977-13982
- Cai DJ, Rickard TC. Reconsidering the role of sleep for motor memory. *Behav Neurosci*, 2009; 123: 1153-1157
- Cohen DA, Pascual-Leone A, Press DZ, Robertson EM. Off-line learning of motor skill memory: A double dissociation of goal and movement. *P Natl Acad Sci USA*, 2005; 102: 18237-18241
- Doyon J, Korman M, Morin A, Dosti V, Tahar AH, Benali H, Karni A, Ungerleider LG, Carrier J. Contribution of night and day sleep vs. simple passage of time to the consolidation of motor sequence and visuomotor adaptation learning. *Exp Brain Res*, 2009; 195: 15-26
- Fischer S, Nitschke MF, Melchert UH, Erdmann C, Born J. Motor memory consolidation in sleep shapes more effective neural representations. *J Neurosci*, 2005; 25: 1248-1255
- Fitts PM. The information capacity of the human motor system in controlling the amplitude of movement. *J Exp Psychol*, 1954; 47: 381-391
- Fogel SM, Smith CT, Cote KA. Dissociable learning-dependent changes in REM and non-REM sleep in declarative and procedural memory systems. *Exp Brain Res*, 2007; 180: 48-61
- Friedman J, Korman M. Kinematic strategies underlying improvement in the acquisition of a sequential finger task with self-generated vs. cued repetition training. *PLoS ONE* 7, 2012; e52063. doi:10.1371

- Hikosaka O, Nakahara K, Rand MK, Sakai K, Lu X, et al. Parallel neural networks for learning sequential procedures. *Trends Neurosci*, 1999; 22: 464-471
- Huber R, Ghilardi MF, Massimini M, Tononi G. Local sleep and learning. *Nature*, 2004; 43: 78-81
- Kempler L, Richmond JL. Effect of sleep on gross motor memory. *Memory*, 2012; 20: 907-914
- Kovacs AJ, Muehlbauer T, Shea CH. The coding and effector transfer of movement sequences. *J Exp Psychol Human*, 2009; 35: 390-407
- Kuriyama K, Stickgold R, Walker MP. Sleep-dependent learning and motor-skill complexity. *Learn Memory*, 2004; 11: 705-713
- Lewis PA, Couch TJ, Walker MP. Keeping time in your sleep: Overnight consolidation in temporal rhythm. *Neuropsychologia*, 2011; 49: 115-123
- Newell A, Rosenbloom PS. Mechanisms of skill acquisition and the law of practice. In Anderson JR (Ed.), *Cognitive skills and their acquisition*. Hillsdale, NJ: Erlbaum, 1-55; 1981
- Newell KM, Mayer-Kress G, Lee Hong SL, Liu YT. Adaptation and learning: Characteristic time scales of performance dynamics. *Hum Movement Sci*, 2009; 28: 655-687
- Panzer S, Krueger M, Muehlbauer T, Shea CH. Asymmetric effector transfer of complex movement sequences. *Hum Movement Sci*, 2010; 29: 62-72
- Pashler H, Johnston JC. Attentional limitations in dual-task performance. In Pashler H (Ed.), *Attention*. Hove, UK: Psychology Press, 155-189; 1998
- Rhodes BJ, Bullock D, Verwey WB, Averbeck BB, Page MPA. Learning and production of movement sequences: Behavioral, neurophysiological, and modelling perspectives. *Hum Movement Sci*, 2004; 23: 699-746
- Robertson EM, Pascual-Leone A, Press DZ. Awareness modifies the skill-learning benefits of sleep. *Curr Biol*, 2004; 14: 208-212
- Sainburg RL. Handedness: Differential specializations for control of trajectory and position. *Exerc Sport Sci Rev*, 2005; 33: 206-213
- Schmidt S, Erlacher D, Blischke K, Brueckner S, Müller F. Sleep related enhancement in motor performance: gross vs. fine motor tasks. *J Sport Exercise Psy*, 2010; 32 (Suppl): 124
- Song S. Consciousness and the consolidation of motor learning. *Behav Brain Res*, 2009; 196: 180-186
- Verwey WB. Buffer loading and chunking in sequential key-pressing. *J Exp Psychol Human*, 1996; 22: 544-562
- Walker MP. A refined model of sleep and the time course of memory formation. *Behav Brain Sci*, 2005; 28: 51-104
- Walker MP, Brakefield T, Morgan A, Hobson JA, Stickgold R. Practice with sleep makes perfect: Sleep-dependent motor skill learning. *Neuron*, 2002; 35: 205-21
- Witt K, Margraf N, Bieber C, Born J, Deuschl G. Sleep consolidates the effector-independent representation of a motor skill. *Neuroscience*, 2010; 171: 227-234
- Wright DL, Rhee J-H, Blischke K, Erlacher D, Brueckner S. Offline improvement occurs for temporal stability but not accuracy following practice of integer and non-integer rhythms. *Acta Psychol*, 2012; 140: 266-273
- Wright DL, Rhee J-H, Vaculin A. Offline improvement during motor sequence learning is not restricted to developing motor chunks. *J Motor Behav*, 2010; 42: 317-324

Corresponding Author:

Klaus Blischke

Saarland University, Institute of Sport Science, P.O. Box 15 11 50

D-66041 Saarbruecken, Germany.

Phone: +49-681-302-4172

E-mail: k.blischke@mx.uni-saarland.de



Chunk Concatenation Evolves with Practice and Sleep-Related Enhancement Consolidation in a Complex Arm Movement Sequence

by

Klaus Blischke¹, Andreas Malangré¹

This paper addresses the notion of chunk concatenation being associated with sleep-related enhancement consolidation of motor sequence memory, thereby essentially contributing to improvements in sequence execution speed. To this end, element movement times of a multi-joint arm movement sequence incorporated in a recent study by Malangré et al. (2014) were reanalyzed. As sequence elements differed with respect to movement distance, element movement times had to be purged from differences solely due to varying trajectory lengths. This was done by dividing each element movement time per subject and trial block by the respective “reference movement time” collected from subjects who had extensively practiced each sequence element in isolation. Any differences in these “relative element movement times” were supposed to reflect element-specific “production costs” imposed solely by the sequence context. Across all subjects non-idiosyncratic, lasting sequence segmentation was shown, and four possible concatenation points (i.e. transition points between successive chunks) within the original arm movement sequence were identified. Based on theoretical suppositions derived from previous work with the discrete sequence production task and the dual processor model (Abrahamse et al., 2013), significantly larger improvements in transition speed occurring at these four concatenation points as compared to the five fastest transition positions within the sequence (associated with mere element execution) were assumed to indicate increased chunk concatenation. As a result, chunk concatenation was shown to proceed during acquisition with physical practice, and, most importantly, to significantly progress some more during retention following a night of sleep, but not during a waking interval.

Key words: motor sequence learning, memory consolidation, offline learning, sleep, gross motor skill.

Introduction

In the motor domain, numerous studies have demonstrated by now that sleep after initial training of motor skills, in the absence of any further physical practice, can produce significant improvements in performance at later retesting. Although in most of these studies fine motor skills like sequential-finger-tapping were used for a criterion task (e.g. Fischer et al., 2005; Walker, 2005; Doyon et al., 2009), recently these findings have also successfully been extended to gross motor tasks involving multi-joint limb

movements of the upper as well as of the lower extremities (Al-Sharman and Siengsukon, 2013; Genzel et al., 2012; Kempler and Richmond, 2012; Malangré et al., 2014; Morita et al., 2012).

This phenomenon, usually referred to as “offline learning” or “enhancement consolidation”, has been associated with an “active system consolidation” process (Born and Wilhelm, 2012). Here it is assumed that newly encoded skill representations are being actively and repeatedly reprocessed during sleep,

¹ - Institute of Sport Science, Saarland University, Saarbruecken, Germany.

resulting in a qualitative reorganization of the respective memory representations (Albouy et al., 2013; Tamminen et al., 2010). Among those qualitative changes in motor sequence memory representation, a prominent role could be ascribed to chunking-processes.

“Chunking” here means parsing long movement sequences into shorter units of integrated motor elements, which can be prepared by a single motor program and executed as a unified action (Verwey, 1996). Since fast and accurate sequence execution requires rapid *in-advance specification* of requisite elements to be organized in immediate succession, motor chunking can essentially facilitate motor sequence performance and effectively reduce the memory load during sequence execution (Bo and Seidler, 2009; Verwey et al., 2015). Preparatory processes associated with the instantiation of a specific sequence chunk typically are reflected in prolonged duration or reaction time of the first as compared to the following elements of that chunk (Verwey, 2001). Theoretically, those differences in element production times are attributed to different *processing phases* within a chunk: chunk selection and preparation are thought to involve higher cognitive processes carried out by the central (or cognitive) processor, while execution of elements comprised within a chunk are triggered by a motor processor from a motor buffer (Abrahamse et al., 2013).

It is assumed that the formation of the respective hierarchical memory structures is driven by two separable processes: one is involved in parsing long strings of contiguous elements into shorter action sets or “chunks” (*segmentation*). The other one involves the transition between adjacent motor chunks (*concatenation*) and is thought to rely on associations between successive motor chunks or on cognitive selection processes. Efficient concatenation is supposed to be essential for the fluid transitions between motor chunks, thus allowing for distinct chunks within a sequence to be executed smoothly and in rapid succession. As a consequence, concatenation points are supposed to appear less and less prominent in the course of sequence skill development. As motor chunks are consolidated, the central processor will still select, load into the motor buffer and initiate each next motor chunk, but will be increasingly relieved

from more detailed processing requirements. With regard to sequences of *aiming* movements such requirements particularly involve specification of movement parameters (i.e. speed, size, absolute force, muscle group, etc.) to each element of the next-following subsequence. After sufficient practice, in addition to the type and number of submovements and movement order, these parameters are also thought to be integrated in the motor chunk representation, rendering a parameter specification stage thereby unnecessary (Verwey et al., 2015). As a result, loading and initiation of such a chunk will proceed faster now.

Motor chunk concatenation usually is associated with extended practice (Sakai et al., 2003), but has also been shown following nocturnal sleep in a sequential finger-tapping task (Kuriyama et al., 2004). Thus, it seems likely that chunk concatenation, in itself a qualitative change in sequence representation, also contributes to quantitative improvements in motor sequence performance (i.e. increased performance speed) typically observed after initial learning at retest following a period of sleep. This notion was tested by reanalyzing behavioral data from a recent study (Malangré et al., 2014), which has proven sleep-related enhancement consolidation in a complex arm movement sequence. Participants initially practiced this task either in the morning (the Morning-Evening-Morning or MEM-group) or in the evening (the Evening-Morning-Evening or EME-group), and were retested 12 hrs (Retest1) and 24 hrs later (Retest2). Thus, subjects were provided with a regular night of sleep either during the first or during the second retention interval. No day-time naps were allowed. *Total Execution Time* per sequence (which is inversely proportional to performance speed) decreased significantly in both experimental groups during acquisition and significantly decreased again in either group during retention *following sleep*, but not following the respective wake intervals (Figure 1). Considering changes in the error rate (i.e. the number of erroneous sequences), there was no indication of any speed-accuracy trade off.

The present paper addresses the following questions: (a) Did *sequence segmentation* occur in the course of the study by Malangré et al. (2014) at all? And if so, (b) did *chunk concatenation* contribute to the sleep-related performance enhancements reported in that study?

To this end, behavioral data from that study were analyzed at the sequence *element* level.

Material and Methods

Subjects

24 subjects (25.5 ± 3.9 years; 6 females, 1 left-handed; 18 males) participated in the original study by Malangré et al. (2014), which was conducted in accordance with the ethical standards of the Declaration of Helsinki (1964). All participants gave their written informed consent. Subjects were randomly assigned to two experimental groups, the Evening-Morning-Evening-group (EME), and the Morning-Evening-Morning-group (MEM), with $N = 12$ each.

Task and procedure in the original study

For a criterion task, subjects originally carried out a 15-element movement sequence with their non-dominant arm. Subjects were seated comfortably in a height-adjustable chair with their upper trunk against the backrest in front of a table-mounted electronic pegboard and a vertical computer screen. With their hand visible all the time, participants could freely move their shoulder, elbow and wrist. On each trial, following a start signal they had to successively fit a small peg into the respective target-holes (depth: 22.22 mm; diameter: 12.7 mm) on the pegboard (Figure 2). The spatial pattern they had to follow was void of any apparent regularity. Transport movements differed in direction and extent (range: 3.83 to 33.75 cm). Precision requirements for sequence elements amounted to an index of difficulty (cf. Fitts, 1954) of 4.95 (± 1.01) on average (Table 1).

Targets were visualized prior to each transport movement by changing the color of the target symbol on the screen from grey to red. Correct execution of a sequence element was indicated by a color change of the respective target stimulus from red to green, while the next target symbol was illuminated red. When a reaching error occurred, the symbol representing the target that had been missed turned green as well, while the next target was illuminated red. Thus, explicit error control always required participants to compare the peg's present position on the pegboard to the target position indicated on the screen. Once a sequence element was terminated, the next movement had to be started immediately, until the sequence was completed.

Start and termination of transport movements were registered by breaking/closing a magnetic contact when the peg was lifted from/inserted into the respective target hole. The respective dwell times (i.e. peg stationary in the hole) proved to be negligible and completely unrelated to the significant treatment effects reported above.

Practice as well as retention tests were organized in successive blocks of 10 trials, separated by a 30 second resting period. Performance measures were the number of Erroneous Sequences per trial block (i.e. the error rate), and Total Execution Time (TET) per sequence, with Total Execution Time averaged for each subject across correct sequences in a trial block. When a reaching error occurred, subjects were to continue the respective sequence trial without correcting for that error. However, suchlike erroneous sequences were excluded from calculation of TET. Initial training amounted to 100 trials (10 blocks of 10 trials), while retests (Retest1 and Retest2) comprised 30 trials (3 blocks of 10 trials) each. Start and Post-Training measures were calculated from the first and the last 30 acquisition trials respectively.

Providing a data base for analyzing performance at the sequence element level

In the original study, Total Execution Time per sequence turned out to be the central outcome variable, with the error rate remaining low throughout the whole experiment and largely unrelated to the treatment. Thus, in the present study, only Total Execution Time data were analyzed, solely reflecting the impact of sleep on performance speed. As a consequence, in order to investigate chunking processes our data reanalysis had to address sequence element duration, which was chiefly expressed by element movement time. However, since reaching movements originally differed in amplitude, elementary movement times within a sequence would have naturally differed due to the different transport distances that had to be covered. Therefore, firstly, discrepancies in element movement times resulting from elements having different trajectory lengths had to be eliminated from the original movement time data.

This was done by relating participants' original sequence element movement times to respective *reference movement times*. The latter were derived from three laboratory staff members

who previously had acquired high expertise in handling the pegboard. These “experts” executed each single sequence element all by itself 20 times in succession as fast as possible with their non-dominant arm. Since each element was carried out *in isolation*, any impact of sequence context on element duration was reliably eliminated. For each element, the ten fastest trials from each expert were averaged. These mean element movement times were highly inter-correlated among experts ($\bar{r} = 0.862$; $p < .001$). They were averaged once more across experts, resulting in element-specific reference movement times, which in turn were significantly correlated with the elements’ index of difficulty ($r = 0.843$; $p < .001$).

Furthermore, for each participant of the original study and each (correct) sequence execution trial, movement time of each single sequence element was determined. These element-specific movement times were then averaged across all correct sequences per subject and trial block. The resulting *mean element movement time* was then divided by the *reference movement time* of that specific element described in the previous paragraph. This resulted in a *relative Element Movement Time* (rEMT) of that sequence element for each subject and trial block. Thus, any *differences* in rEMT values still present within a subject’s sequence trial are supposed to reflect element-specific “production costs” imposed solely by an element’s *position within the sequence*.

Assessing chunk concatenation

Chunk concatenation was assumed whenever rEMTs of putative concatenation points (i.e. *long* rEMTs) were shown to decrease *relatively more* than rEMTs of elements classified as integral parts of a chunk (i.e. *short* rEMTs), thereby representing the mere execution phase. Thus, for each group and time point interval (Acquisition, Retention 1, Retention 2), it was tested if rEMTs of putative concatenation points decreased significantly more than rEMTs of elements identified as integral parts of chunks.

Particularly long and short rEMTs characterizing either putative concatenation points or integral parts of chunks were determined by a comprehensive comparison of all element-specific rEMT values, as explained in detail in the results section below (see the paragraph on sequence segmentation).

Results

For inferential statistics, two- and three-way ANOVAs were run. With respect to repeated-measures factors, in case of violation of the sphericity assumption *df*-correction according to Greenhouse-Geisser was applied. A significance level of $p \leq .05$ was used for all inferential statistics. Calculations were conducted with SPSS-PC, version 15.0.

Error rate

In the original study, over all error rate was small right from the beginning (i.e. 2.8 ± 2.01 Erroneous Sequences per trial block on average), and did not fluctuate significantly across trial blocks during acquisition ($p = .079$). There was no Group \times Block interaction ($p = .273$), nor did groups differ in performance ($p = .374$). During retention, again the error rate did not change across time points ($p = .239$), nor did groups differ ($p = .904$). Thus, the number of Erroneous Sequences remained largely unaffected by treatment conditions during retention (i.e. participants either sleeping or staying awake), and did not increase as execution speed increased. There were also no indications for reaching errors to systematically occur at preferred locations within the criterion sequence. Thus, the error rate did not provide any indication of sequence segmentation.

Sequence segmentation

According to a 2[Group] \times 15[Element] \times 16[Block] ANOVA (with rep. measures on the last two factors), rEMTs differed and decreased significantly across blocks ($F_{[\text{Element}]}$ (5.629, 123.836) = 47.379, $p < .001$, $\eta^2_p = .683$; $F_{[\text{Block}]}$ (2.972, 65.391) = 59.508, $p < .001$, $\eta^2_p = .730$). The rEMT-patterns in both groups were not different, since neither the “Group” main effect ($p = .362$), nor the “Element \times Group” ($p = .374$), nor the triple-interaction ($p = .708$) reached statistical significance.

Presence of sequence segmentation then was demonstrated by a comprehensive comparison of all element-specific rEMT values. The respective mean rEMTs for each sequence element (collapsed across all subjects and all 16 trial blocks) with their respective 95%-confidence intervals are listed in Table 1. Two mutually exclusive sets of 5 particularly *long* and 6 particularly *short* rEMTs were identified (Duncan Post-Hoc-Test). The “long” rEMT set comprised sequence elements 4,

6, 7, 11 and 13. The “short” rEMT set included elements 1, 3, 5, 8, 9 and 15. rEMTs of elements 2, 10, 12 and 14 were considered “intermediate”. Over all rEMTs were not related to the elements’ index of difficulty any more ($r = -.440$; $p = .101$).

Non-idiosyncratic segmentation patterns

To obtain a measure for segmentation similarity across members of an experimental group, we computed the average rEMT pattern in our 15-element arm movement sequence at each time point (i.e. Start, Post-Training, Retest1, and Retest2) for each group. These group-average patterns were used as reference, and at each time point correlations (Pearson’s r) were computed for the rEMT-pattern of each subject with the respective group average (cf. Verwey and Eikelboom, 2003). A 2[Group] \times 4[Time Point] ANOVA on those Fisher Z-transformed correlations with repeated measures on the factor “Time Point” showed that groups did not differ ($p_{\text{[Group]}} = .701$; $p_{\text{[Group} \times \text{Time-Point]}} = .530$), but revealed a significant “Time Point” main effect ($F(3, 66) = 10.114$, $p < .001$, $\eta^2_p = .315$). Following respective pairwise comparisons, similarity of individual rEMT patterns with the respective group-average in our task decreased somewhat with practice (Start – Post-Training: $F(1, 22) = 25.131$, $p < .001$, $\eta^2_p = .533$), but then remained stable across retention ($p \geq .469$). As indicated by the fairly high correlation coefficients (Pearson’s r of individual rEMT patterns with group-average patterns, averaged across all 24 subjects: Start: .873; Post-Training: .737; Retest1: .757; Retest2: .775), different participants had altogether very similar segmentation patterns at each time point throughout the whole experiment.

Functional classification of rEMT-clusters

Drawing on the notion of different processing phases, the rEMT clusters identified above were functionally characterized as follows: (a) Particularly “long” rEMTs immediately followed by at least one element characterized by a “short” or an “intermediate” rEMT most likely indicate typical chunk *concatenation points* (i.e. elements 4, 7, 11 and 13). (b) Particularly “short” rEMTs immediately following assumed concatenation points are thought to reflect the mere *execution* process (i.e. elements 3, 5, 8, 9 and 15). (c) *Intermediate* rEMT values (elements 2, 10, 12 and 14) could possibly be caused by individual differences, but could also designate elements not

fully integrated into a motor chunk yet because of sequence knowledge still being somewhat weak.

Special cases: element 6 cannot be considered a concatenation point, because it was nearly always followed by another “long” rEMT (i.e. element 7). Element 1 represents sequence initiation, usually associated with a comparatively long preparation time. In the original study, however, participants were free to initiate the criterion sequence within a certain time limit following the start-signal. This preparation time preceding actual sequence initiation at that time had not been recorded and was not included in the TET measure. Therefore, in the present data-reanalysis sequence initiation was not taken into account.

Chunk concatenation

Based on the two distinct sets of four “long” rEMTs considered putative concatenation points (i.e. elements 4, 7, 11 and 13), and five “short” rEMTs considered integral parts of chunks (i.e. elements 3, 5, 8, 9 and 15), the factor “Processing Phase” (levels: “concatenation” = long rEMTs, “execution” = short rEMTs) was subjected to further analysis. Three 2[Group] \times 2[Time Point] \times 2[Processing Phase] ANOVAs (rep. measures on the last two factors) were conducted separately for acquisition and each retention interval (Acquisition: Start – Post-Training; Retention 1: Post-Training – Retest1; Retention 2: Retest1 – Retest2).

Globally, long rEMTs decreased significantly more than short ones during the first two intervals, but not during the third any more (Acquisition: $F_{\text{[Time Point} \times \text{Processing Phase]}}(1, 22) = 30.173$, $p < .001$, $\eta^2_p = .578$; Retention 1: $F_{\text{[Time Point} \times \text{Processing Phase]}}(1, 22) = 7.417$, $p = .012$, $\eta^2_p = .254$; Retention 2: $F_{\text{[Time Point} \times \text{Processing Phase]}}(1, 22) = 2.504$, $p = .128$). While groups did not differ in this respect during acquisition (triple interaction: $F(1, 22) = .211$, $p = .650$), the respective *triple interactions* turned out *significant* for each of the two subsequent *retention intervals* (Retention 1: $F(1, 22) = 7.900$, $p = .010$, $\eta^2_p = .264$; Retention 2: $F(1, 22) = 5.931$, $p = .023$, $\eta^2_p = .212$). That is, across each of the retention intervals the two experimental groups exhibited significantly different patterns of change in rEMTs (Figure 3).

According to our expectation of sleep boosting chunk concatenation, this should have been the case indeed if long rEMTs had decreased more than short ones in subjects provided with sleep

during the respective retention interval than in those subjects that had to stay awake.

At this point, for final clarification detailed follow-up analyses were required. With respect to the repeated measures factors “Time Point” and “Processing Phase” (levels: long rEMTs, short rEMTs), 2[Time Point] x 2[Processing Phase] repeated measures-ANOVAs were calculated separately for each group and each retention interval (Retention 1: Post-Training – Retest1; Retention 2: Retest1 – Retest2), respectively. For both groups in both retention intervals long and short rEMTs significantly differed ($p < .001$). During *Retention 1*, there was no significant change of rEMTs in the MEM group any more ($F_{[\text{Time Point}]}$ (1, 11) = 4.581, $p = .056$, $\eta^2_p = .294$; interaction: $p = .961$), while in the EME group (with sleep provided during this first interval) the main effect “Time Point” as well as the interaction

term turned out significant ($F_{[\text{Time Point}]}$ (1, 11) = 37.903, $p < .001$, $\eta^2_p = .775$; $F_{[\text{Time Point} \times \text{Processing Phase}]}$ (1, 11) = 19.321, $p = .001$, $\eta^2_p = .637$).

During *Retention 2*, however, the main effect “Time Point” and the interaction term turned out significant this time in the MEM group, i.e. the group provided with sleep during this second retention interval ($F_{[\text{Time Point}]}$ (1, 11) = 30.566, $p < .001$, $\eta^2_p = .735$; $F_{[\text{Time Point} \times \text{Processing Phase}]}$ (1, 11) = 5.327, $p = .041$, $\eta^2_p = .326$), while in the EME group neither rEMT changes nor interactions were significant any more ($p \geq .405$).

Thus, comparison of changes over time between the two sets of rEMTs revealed that in each experimental group the *long* rEMTs *decreased significantly more* than short ones during *practice*, and then once more during *retention* when tested *following sleep*, but not when tested *following wake*.

Table 1

Relative Element Movement Time (rEMT) and Index of Difficulty (ID).

Element	ID	rEMT		
		short	intermediate	long
1	5.74	1.79 (1.70 – 1.89)		
2	4.59		1.97 (1.84 – 2.11)	
3	5.21	1.81 (1.70 – 1.92)		
4	2.58			2.24 (2.05 – 2.42)
5	5.01	1.85 (1.76 – 1.95)		
6	5.18			2.35 (2.24 – 2.46)
7	5.11			2.71 (2.55 – 2.87)
8	5.60	1.70 (1.62 – 1.78)		
9	5.86	1.84 (1.76 – 1.91)		
10	5.46		2.09 (1.98 – 2.21)	
11	2.58			2.38 (2.15 – 2.61)
12	5.01		2.07 (1.95 – 2.20)	
13	5.46			2.28 (2.15 – 2.41)
14	5.11		1.98 (1.87 – 2.08)	
15	5.74	1.80 (1.71 – 1.90)		

Annotations: Reported are for each sequence element (a) Fitts' index of difficulty (ID = $\log_2 [2A/W]$; A = amplitude, and W = target width), and (b) rEMT mean values (dimensionless) for each sequence element, calculated across all 24 subjects and all 16 trial blocks, and the respective 95% confidence intervals (in parentheses). Note that upper confidence interval limits of short rEMTs and lower confidence interval limits of long rEMTs do not overlap. See also the results section for classification of rEMT values.

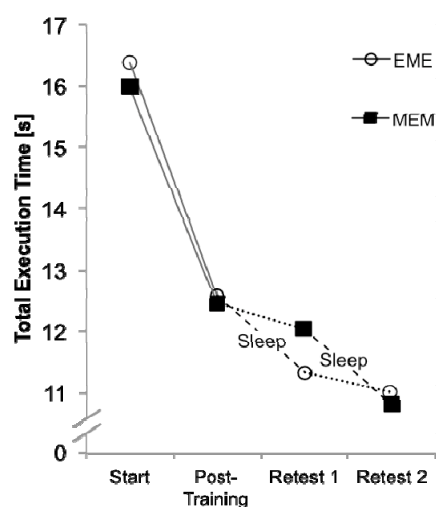


Figure 1

Mean total sequence execution time (TET [s]; correct sequences only) per group (MEM, EME) and time point (Start: block 1, 2 & 3; Post-Training: block 8, 9 & 10; Retest 1 (after 12 hrs): block 11, 12 & 13; Retest 2 (after 24 hrs): block 14, 15 & 16). Acquisition: solid lines. Retention: dotted lines = wake periods; dashed lines = sleep-filled periods (data from Malangré et al., 2014).

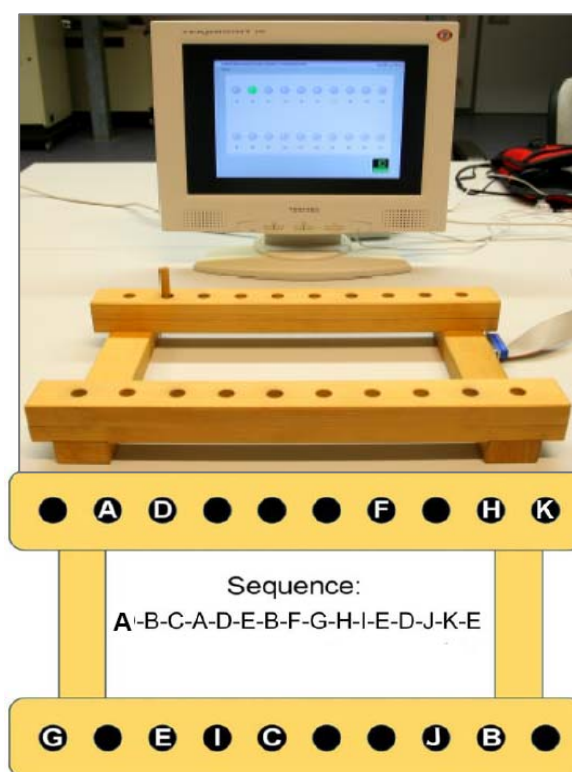


Figure 2

Experimental apparatus (upper panel), and spatial locations to be reached for one after the other, defining the fifteen-element arm movement sequence (lower panel).

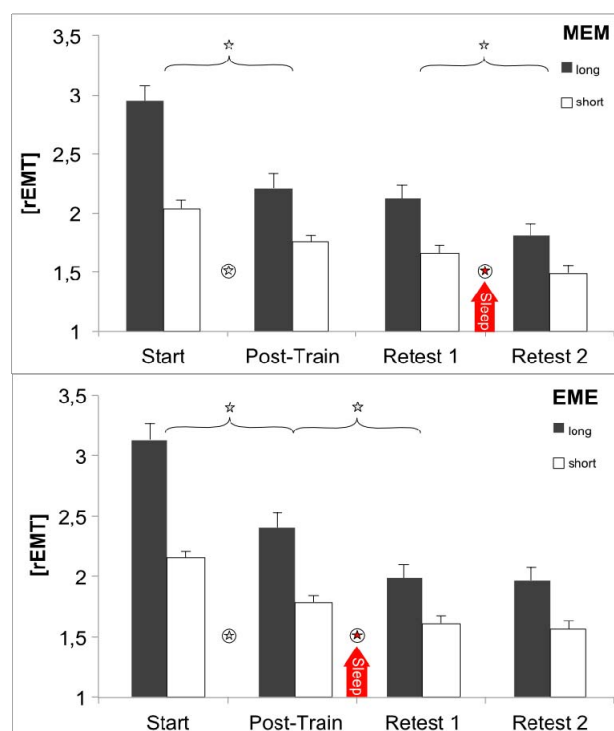


Figure 3

Changes in relative Element Movement Time (rEMT) representing (a) the concatenation level (filled bars: “long” rEMTs; elements 4, 7, 11 and 13) and (b) the execution level (open bars: “short” rEMTs; elements 3, 5, 8, 9 and 15). Presented are group means per time point. Error bars: standard errors of the mean. Top panel: MEM group; bottom panel: EME group. For each time point interval, simple stars indicate a significant main effect “Time Point”; circled stars stand for a significant interaction “Time Point” x “Processing Phase” [i.e. rEMT level].

Discussion

In order to better understand what kind of representational changes take place in the course of an “active system consolidation” process in motor skills during sleep, performance data from a recent study by Malangré et al. (2014) were reanalyzed. In that study, a sequence of 15 precise reaching movements had to be executed as fast as possible with as few errors as possible. For the purpose of data reanalysis, participants’ original element-specific movement times were subjected to a transformation procedure in order to eliminate differences in movement duration resulting purely from dissimilar transport distances which had to be covered for different

sequence elements. The resulting dimensionless “relative Element Movement Times” (rEMTs) eventually allowed participants’ performance to be addressed at the sequence element level. As it was then evidenced by the present data reanalysis, sleep-related offline improvements in performance speed found in the original study were closely associated with increased concatenation of adjacent sets of sequence elements (i.e. chunks).

Sequence segmentation is found to be largely non-idiosyncratic

Presence of *sequence segmentation* then was demonstrated by comparison of element-specific rEMT values. Three different sets of rEMTs

were identified: five were long, six were short and four were intermediate. Drawing on the concept of different processing phases in sequence production (i.e. initiation, concatenation and execution) proposed in the context of the dual processor model (Abrahamse et al., 2013; Verwey et al., 2015), those elements with particularly *long* rEMTs *immediately preceding* short or intermediate ones were understood as possible chunk *concatenation* points, indexing the transition from one motor chunk to the next (the concatenation phase). The resulting segmentation pattern turned out *fairly consistent* across time points and appeared also to be largely *non-idiosyncratic*.

This requires closer consideration for it has been shown previously participants sometimes tend to develop *individually different* segmentation patterns when rapidly executing key-pressing sequences with little or no apparent regularities (Kuriyama et al., 2004; Sakai et al., 2003). More specifically, Verwey and Eikelboom (2003) showed that the amount of idiosyncratic sequence segmentation differed for different sequences and was virtually identical for all subjects only with one particular sequence. Thus, in order to check if our results were possibly obscured by initially averaging rEMTs across subjects and trial blocks, we also applied *k-means clustering* (MacQueen, 1967) for motor chunk identification (cf. Song and Cohen, 2014). This data-driven procedure does not require any a priori assumptions on chunk length or on the ratio of long and short rEMTs. It is sensitive to idiosyncratic chunking behavior (which means the concatenation points are found to occur at different sequential locations for different people) and it allows tracing chunk development across time points. As reported in Appendix 1, our above finding of sequence segmentation turning out *non-idiosyncratic* and fairly stable across retention was fully confirmed.

Sequence segmentation being non-idiosyncratic in our study is explained most parsimoniously by specific task constraints which impose segmentation at the same locations within the sequence across participants. In finger-tapping tasks, sequence segmentation was imposed artificially by (a) introducing regularities in response order, (b) temporarily inserting a pause between two successive stimuli, (c) using differently colored key-specific stimuli to

distinguish subsequences (Abrahamse et al., 2013). None of these conditions immediately applies to our arm movement sequence. However, reliably locating the next target on the pegboard always proved to be difficult, when there was a whole string of holes extending to each side of the new target and long transport distances had to be covered. To execute such elements correctly, early on in practice required elaborate visual search and implementation of auxiliary strategies (e.g. counting out holes). This in turn could have had an effect equivalent to imposing an artificial pause at this point of the sequence. The same holds true where the “dynamical flow” of sequence production was interrupted on account of the spatial arrangement of two successive target holes.

Thus, all in all our reanalysis revealed a (none-idiosyncratic) segmentation pattern of five sub-sets (including the one immediately following sequence initiation) comprising two to four elements each. This result nicely reconciles the notion that longer, fixed movement sequences (> 4 elements) generally are subdivided into a succession of multiple motor chunks (Abrahamse et al., 2013; Bo and Seidler, 2009). The number and size of sub-sets found in our 15-element arm movement sequence are well in accordance with capacity limitations of working memory (Cowan, 2010) as well as the notion of the motor buffer being limited to about 3 to 5 elements, thereby defining the average length of a single motor chunk (Bo and Seidler, 2009; Verwey and Eikelboom, 2003).

Chunk concatenation evolves with sleep and physical practice

In the discrete sequence production task, sequence segmentation was shown to be maintained even after considerable practice (Verwey and Eikelboom, 2003). Lasting segmentation might imply that the length of motor chunks is structurally limited, although there may be individual differences as to how many elements are comprised into one chunk. This considered, we did not expect the absolute *number* of concatenation points initially identified in our arm movement sequence to significantly decrease across time points on account of smaller chunks being totally concatenated into longer ones. As it turned out, participants' segmentation patterns were indeed well preserved throughout

the whole experiment, and the number of concatenation points on average did not change.

We did, however, expect relative Element Movement Times at chunk concatenation points to *decrease significantly more* during a sleep-filled retention period than relative Element Movement Times reflecting merely rapid execution *within* a chunk. This notion of a selective sleep-related offline improvement had already been proposed by Kuriyama et al. (2004) more than ten years ago. These authors found sleep-related performance enhancements (increased execution speed and error reduction) in four finger-tapping tasks of different length (five-elements or nine-elements) and complexity (unimanual or bimanual). They also assessed each participant's slowest ("problem point") and fastest ("easy point") transition from one key press to the next in each sequence. At the group level, they found significant overnight improvements in transition speed occurring at the slowest, but not at the fastest transition position. This selective improvement was specific to the overnight learning process and was neither found during acquisition, nor during a wake retention interval in a control group.

Other than Kuriyama et al. (2004), we did not compare just two transition points between sequence elements. Instead we compared two whole sets of elements which were either classified as putative concatenation points (long rEMTs) or identified as integral parts of a chunk (short rEMTs) before. During *retention*, in each of our experimental groups those *long* rEMTs decreased *significantly more* than the short ones *following* sleep, while there were no significant changes in rEMTs following wake. According to the "Cognitive framework for Sequential Motor Behavior" (Verwey et al., 2015), this selective improvement may be conceived as the central processor becoming more efficient and thereby faster at selecting, retrieving and loading each chunk into the motor buffer for subsequent execution during sleep, but not during wake.

Interestingly, we observed the same selective improvement also during *acquisition*. Here, both sets of rEMTs significantly decreased, but the *long* ones again decreased *significantly more* than the short ones. While Kuriyama et al. (2004) also found transition points to improve at a different rate across acquisition, they notwithstanding interpreted significant

improvements of *both* transitions (fastest *and* slowest) in evidence of chunk concatenation being completely *absent* during initial training. In the light of our present findings, this position ought to be reconsidered. We would rather argue that progressive chunk concatenation always comes about as an integral part of sequence learning, either "online" associated with overt physical activity and processing of error information or "offline" associated with a sleep-dependent rearrangement of the respective movement representation (i.e. enhancement consolidation).

These findings also bear implications relevant to sports training. This would specifically pertain to complex motor skills requiring fast and precise execution of a series of sub-movements in Euclidian space, like, e.g., performing a karate kata or isolation dance routine. Memory representations of suchlike skills are almost certainly subject to chunking processes. Efficient chunk concatenation will make reproduction of such a skill easier and more reliable specifically under free recall conditions in a tournament. Thus, whenever motor tasks of this kind have to be learned or relearned efficiently, we would recommend practice to be distributed over two or more sessions separated by night or daytime sleep respectively in order to facilitate chunk formation and concatenation. Distributed practice schedules of this kind might be especially beneficial in situations, where the amount of practice per session is limited of necessity, as is often the case in rehab-training or acquisition of sport skills requiring extraordinary high physical or mental effort.

Conclusion

In this paper, performance data of a recent study by Malangré et al. (2014) showing sleep-related offline learning in a complex arm movement sequence were reanalyzed at the sequence element level. This way motor chunk concatenation was shown to evolve in the course of sleep-related enhancement consolidation, thereby making up a relevant portion of overnight improvements in sequence execution speed reported previously. Chunk concatenation also proceeded with physical practice, but was not observed at retests following a waking retention interval. To our knowledge for the first time, such qualitative changes in sequence memory

representation have been shown in a task involving unrestrained multi-joint movements of an entire limb. These findings corroborate and complement earlier findings by Kuriyama et al. (2004), and successfully extend the respective theoretical suppositions based on previous work with the Discrete Sequence Production-task and derived from the Dual Processor Model and the “Cognitive framework for Sequential Motor Behavior” to gross motor abilities closer to so

many every-day and vocational activities. Occurrence of non-idiosyncratic sequence segmentation was plausibly attributed to constraints imposed by features specific to our criterion task. This might encourage researchers to increasingly consider task-specific constraints imposed by tool use and environment, and thereby to arrive at conclusions of improved ecological validity.

Acknowledgements

Part of this work was presented (a) in a poster presentation at the annual conference of the North American Society for the Psychology of Sport and Physical Activity (NASPSPA), June 12th-14th, 2014, at Minneapolis, MN, USA (Abstract published in conference proceedings: Blischke K, Malangré A. Chunk concatenation essentially contributes to sleep-related enhancement consolidation in a complex arm movement sequence. *J Sport Exercise Psy*, 2014; 36 (Suppl): 20, (b) in an oral presentation at the 19th annual conference of the European College of Sport Science (ECSS), July 2nd-5th, 2014, at Amsterdam, Netherlands (Malangré A, Blischke K. Sleep boosts chunk concatenation in a gross motor task.), and (c) in an oral presentation at the 14th bi-annual conference of the dvs-division “Sportmotorik”, March 5th-7th, 2015, at Munich, Germany (Abstract published in conference proceedings: Blischke K, Malangré A. Sleep enhances chunk concatenation in a complex arm movement sequence. In Hermsdörfer J, Stadler W, Johannsen L (eds.). *The Athlete's Brain: Neuronale Aspekte motorischer Kontrolle*. Hamburg: Feldhaus – Edition Czwilina, 53-54; 2015).

We are grateful to an anonymous reviewer for his well informed and constructive comments on an earlier version of this paper, and we also thank F. Döhring for assistance with k-means clustering.

Appendix 1

K-means clustering. First, for each participant element-specific rEMT values were averaged across trial blocks per time-point (Start, Post-Training, Retest1 and Retest2). Then, the clustering procedure (matlab ‘kmeans’, cluster number = 2, empty action = drop, outliers beyond 2 SD of the mean rEMT are set to 2SD of the mean) was applied on each individual’s rEMT-data set per time point. With a cluster number = 2, rEMTs of the 15 sequence elements self-sorted as either long (i.e. slow element) or short (i.e. fast element).

Sequence segmentation. After clustering, motor chunks were easily identified as one or more short rEMTs (i.e. fast elements) following sequence initiation or one of the long rEMTs. With the first (sequence initiation) and the last element excluded, 261 such possible individual concatenation points could be identified in all 24 subjects across the four time points (Start, Post-Training, Retest1 and Retest2). Of these a total of 204 (78.2%) were allotted to the same four sequence elements, namely the elements 4 (16.8%), 7 (35%), 11 (9.6%) and 13 (16.8%). The rest (21.8%) was distributed over eight other sequence elements ($\leq 5.4\%$ per element). Including the initial subset of elements following sequence initiation, subjects on the average subdivided the 15-element movement sequence into $3.61 \pm .96$ chunks (SE: $\pm .11$).

Possible differences in the number of chunks between groups and across time points (TP) were analyzed by means of a 2[Group] \times 4[Time Point] ANOVA (rep. measures on “Time Point”). This ANOVA altogether failed to reach significance ($F_{[\text{Time Point}]}$ (3, 66) = 2.050, $p = .115$; $F_{[\text{Group}]}$ (1, 22) = .223, $p = .642$; no significant interaction: F (3, 66) = 1.162, $p = .331$). Once established during acquisition, the resulting pattern of slow and fast elements appeared to be fairly stable across both retention intervals, as indicated by highly significant correlations ($p < .01$) of segmentation patterns between the respective time points (mean correlation coefficients calculated from Fisher’s Z-transformed correlations at the subject level; Pearson’s r : .782 (Post-Training – Retest1), and .682 (Retest1 – Retest2)).

Total chunk concatenation. According to Song and Cohen (2014), as smaller chunks are *completely* (!) concatenated into longer chunks, the ratio of the *maximum* number of “fast” elements comprised in one

chunk to the *total* number of “fast” elements within *all* motor chunks should increase and eventually approach one (fully concatenated). This ratio was calculated for each participant at each time point. On these ratios, a 2[Group] x 4[Time Point] ANOVA (rep. measures on “Time Point”) was run. Neither the main effects ($F_{[\text{Time Point}]}$ (3, 66) = 1.825, p = .151; $F_{[\text{Group}]}$ (1, 22) = .379, p = .544), nor the interaction turned out significant (F (3, 66) = .4402, p = .725). Thus, at the group level no indication of *total* chunk concatenation across time points was found.

References

- Abrahamse EL, Ruitenberg MFL, de Kleine E, Verwey WB. Control of automated behaviour: insights from the discrete sequence production task. *Front Hum Neurosci*, 2013; 7: Article 82
- Al-Sharman A, Siengsukon CF. Sleep enhances learning of a functional motor task in young adults. *Phys Ther*, 2013; 93: 1625-1635
- Albouy G, King BR, Maquet P, Doyon J. Hippocampus and striatum: Dynamics and interaction during acquisition and sleep-related motor sequence memory consolidation. *Hippocampus*, 2013; 23: 985-1004
- Blischke K, Erlacher D, Kresin H, Brueckner S, Malangré A. Benefits of sleep in motor learning – prospects and limitations. *Journal of Human Kinetics*, 2008; 20: 23-35
- Bo J, Seidler RD. Visuospatial working memory capacity predicts the organization of acquired explicit motor sequences. *J Neurophysiol*, 2009; 101: 3116-3125
- Born J, Wilhelm I. System consolidation of memory during sleep. *Psychol Res*, 2012; 76: 192-203
- Cowan N. The magical mystery four: How is working memory capacity limited, and why? *Curr Dir Psychol Sci*, 2010; 19: 51-57
- Doyon J, Korman M, Morin A, Dosti V, Tahar AH, Benali H, Karni A, Ungerleider LG, Carrier J. Contribution of night and day sleep vs. simple passage of time to the consolidation of motor sequence and visuomotor adaptation learning. *Exp Brain Res*, 2009; 195: 15-26
- Fischer S, Nitschke MF, Melchert UH, Erdmann C, Born J. Motor memory consolidation in sleep shapes more effective neural representations. *J Neurosci*, 2005; 25: 1248-1255
- Fitts PM. The information capacity of the human motor system in controlling the amplitude of movement. *J Exp Psychol*, 1954; 47: 381-391
- Genzel L, Quack A, Jäger E, Konrad B, Steiger A, Dresler M. Complex motor sequence skills profit from sleep. *Neuropsychobiology*, 2012; 66: 237-243
- Hoedlmoser K, Birklbauer J, Schabus M, Eibenberger P, Rigler S, Mueller E. The impact of diurnal sleep on the consolidation of a complex gross motor adaptation task. *J Sleep Res*, 2015; 24: 100-109
- Kempler L, Richmond JL. Effect of sleep on gross motor memory. *Memory*, 2012; 20: 907-914
- Kuriyama K, Stickgold R, Walker MP. Sleep-dependent learning and motor-skill complexity. *Learn Memory*, 2004; 11: 705-713
- MacQueen J. Some methods for classification and analysis of multivariate observations. In Lecam L, Neyman J (eds.). *Proc 5th Berkely Symp Math Stat Prob 1965/66*. Berkely, 1: 281-297; 1967
- Malangré A, Leinen P, Blischke K. Sleep-related offline learning in a complex arm movement sequence. *Journal of Human Kinetics*, 2014; 40: 7-20
- Morita Y, Ogawa K, Uchida S. The effect of a daytime 2-hour nap on complex motor skill learning. *Sleep and Biological Rhythms*, 2012; 10: 302-309
- Sakai K, Kitaguchi K, Hikosaka O. Chunking during human visuomotor sequence learning. *Exp Brain Res*, 2003; 152: 229-242
- Song S, Cohen L. Impact of conscious intent on chunking during motor learning. *Learn Mem*, 2014; 21: 449-451
- Tamminen J, Payne JD, Stickgold R, Wamsley EJ, Gaskell MG. Sleep spindle activity is associated with the integration of new memories and existing knowledge. *J Neurosci*, 2010; 30: 14356-14360
- Verwey WB. Buffer loading and chunking in sequential key-pressing. *J Exp Psychol Human*, 1996; 22: 544-562
- Verwey WB. Concatenating familiar movement sequences: The versatile cognitive processor. *Acta Psychol*, 2001; 106(1-2): 69-95

- Verwey WB, Eikelboom T. Evidence for lasting sequence segmentation in the discrete sequence-production task. *J Mot Behav*, 2003; 35: 171-181
- Verwey WB, Shea CH, Wright DL. A cognitive framework for explaining serial processing and sequence execution strategies. *Psychon Bull Rev*, 2015; 22: 54-77
- Walker MP. A refined model of sleep and the time course of memory formation. *Behav Brain Sci*, 2005; 28: 51-104

Corresponding author:

Klaus Blischke

Saarland University, Institute of Sport Science
Campus B8-2
D-66123 Saarbruecken, Germany.
Phone: +49-681-302-4172
e-mail: k.blischke@mx.uni-saarland.de



Sleep-Related Offline Improvements in Gross Motor Task Performance Occur Under Free Recall Requirements

Andreas Malangré and Klaus Blischke*

Sport Science, Training Science, Saarland University, Saarbruecken, Saarland, Germany

Nocturnal sleep effects on memory consolidation following gross motor sequence learning were examined using a complex arm movement task. This task required participants to produce non-regular spatial patterns in the horizontal plane by successively fitting a small peg into different target-holes on an electronic pegboard. The respective reaching movements typically differed in amplitude and direction. Targets were visualized prior to each transport movement on a computer screen. With this task we tested 18 subjects (22.6 ± 1.9 years; 8 female) using a between-subjects design. Participants initially learned a 10-element arm movement sequence either in the morning or in the evening. Performance was retested under free recall requirements 15 min post training, as well as 12 and 24 h later. Thus, each group was provided with one sleep-filled and one wake retention interval. Dependent variables were error rate (number of Erroneous Sequences, ES) and average sequence execution time (correct sequences only). Performance improved during acquisition. Error rate remained stable across retention. Sequence execution time (inverse to execution speed) significantly decreased again during the sleep-filled retention intervals, but remained stable during the respective wake intervals. These results corroborate recent findings on sleep-related enhancement consolidation in ecological valid, complex gross motor tasks. At the same time, they suggest this effect to be truly memory-based and independent from repeated access to extrinsic sequence information during retests.

Keywords: sleep, enhancement consolidation, gross motor task, sequence learning, free recall

OPEN ACCESS

Edited by:

Genevieve Albouy,
KU Leuven, Belgium

Reviewed by:

Bjoern Rasch,
University of Fribourg, Switzerland
Arnaud Boutin,
University of Montreal, Canada

*Correspondence:

Klaus Blischke
k.blischke@mx.uni-saarland.de

Received: 30 October 2015

Accepted: 14 March 2016

Published: 29 March 2016

Citation:

Malangré A and Blischke K (2016)
Sleep-Related Offline Improvements
in Gross Motor Task Performance
Occur Under Free Recall
Requirements.
Front. Hum. Neurosci. 10:134.
doi: 10.3389/fnhum.2016.00134

INTRODUCTION

There is ample evidence by now that sleep (but not wake) after initial training of motor skills can produce significant improvements in performance at later retesting without any further physical practice (e.g., Fischer et al., 2005; Walker, 2005; Doyon et al., 2009). This phenomenon usually is referred to as “sleep-related offline learning”, and has been associated with an “active system consolidation” process (Born and Wilhelm, 2012). Here, it is assumed that newly encoded skill representations are being actively reprocessed during slow-wave sleep, resulting in strengthening synaptic connections in the neocortex and in a qualitative reorganization of the respective memory representations. These processes are understood as a prerequisite for the sudden improvements in overt performance frequently observed. However, suchlike processes and effects appear to be closely related to certain motor task characteristics as well as to

specific learning procedures. That is, in general sleep-related offline learning seems to require some involvement of *declarative* memory processes. This is often associated with routines of *explicit* learning and awareness (Robertson et al., 2004). Enhancement of motor sequence memory supposedly pertains to an abstract spatial map of the sequence that represents the series of movements to perform during recall. This representation is supported by a distinct hippocampo-cortical neural network (Albouy et al., 2015), and is supposedly associated with declarative knowledge concerning the action's goal as well as the type of sequence elements and their temporal order. Moreover, performance improvements in motor adaptation tasks (i.e., precise sub-maximal force production; visuo-motor adaptation) have been found to be fairly small and rather *time*-instead of sleep-dependent (Blischke et al., 2008; Doyon et al., 2009; but see Huber et al., 2004). Thus, sleep-related EC should be most pronounced in *sequentially* structured motor tasks.

In most of the studies addressing sleep-related motor offline learning sequential-finger-tapping or thumb-to-finger movements were involved. Only a couple of years ago the question has been raised if the respective findings also apply to ecologically valid gross motor skills (Blischke et al., 2008). And it was only recently that these findings have successfully been extended to gross motor tasks involving multi-joint limb movements (Genzel et al., 2012; Kempler and Richmond, 2012; Morita et al., 2012; Al-Sharman and Siengsukon, 2013; Malangré et al., 2014). Moreover it has been shown that the degree of sleep-related motor enhancement consolidation in the elderly is modulated by the kinematic demands of the task. In one recent study, sleep-related performance improvements were observed in older age groups only when a classic sequence learning task requiring individuated finger movements was replaced by an adapted version of the same task. In this adapted version reaching movements were performed with the whole hand (Gudberg et al., 2015). This dissociation of specific mechanisms of sleep underpinning motor sequence consolidation in older adults is certainly of theoretical importance. And it emphasizes the potential of incorporating whole limb movements in research activities concerning the relation of sleep and motor memory consolidation.

Although criterion tasks incorporated in all these studies reporting gross motor sleep-related offline learning were of considerable variety and involved movements of the upper as well as of the lower extremities, again they were all *sequentially* structured. However, when motor *adaptation* was the prominent task requirement, sleep did not enhance, but only stabilize performance (Hoedlmoser et al., 2015). Thus, the above mentioned dissociation of motor sequence learning and motor adaptation with respect to sleep-related memory consolidation processes seems to hold also for gross motor skills. However, there are still some aspects of practical and theoretical importance waiting for closer scrutiny. One such aspect is the question as to whether sleep-related offline learning will also come into effect at retention even under *free recall* conditions at an early learning stage. This question is of particular importance in the applied field of movement studies

(i.e., vocational training, sports, occupational therapy, and motor rehabilitation). Here trainees, athletes and patients initially are supplied with stimulus information and feedback while acquiring new motor skills at initial training sessions. But soon after initial training they are usually required to recall and execute those skills under “real-life” conditions in the absence of any augmented information.

Here as a first step we present an experiment set up to scrutinize if sleep-related offline learning was to be found at all in a gross motor task under *free recall* conditions with no extrinsic criterion information available. The criterion movement employed was a sequential motor task with high demands on precision and manual dexterity. This task incorporated a series of 10 unrestrained multi-joint reaching movements involving the whole non-dominant arm. Such a task bears good resemblance to a wide variety of sport skills and activities of daily living. Following a fixed spatial pattern, participants had to execute this movement sequence as rapidly as possible with as few errors as possible.

It was hypothesized that after initial learning *sleep*, but not *wake*, significantly facilitates performance (namely: sequence execution speed) at retention under free recall conditions when compared to post-training performance (i.e., free-recall performance assessed shortly after acquisition).

MATERIALS AND METHODS

Participants and Groups

Two groups of participants ($N = 12$ each) voluntarily participated in this experiment, which was conducted at the Saarland University (Department of Sport Science) in accordance with the ethical standards of the 1975 Declaration of Helsinki, and was approved by the Ethic Committee of the Faculty 5 Empirical Social Sciences of Saarland University. Subjects took part in the experiment in accord with the department's course regulations and gave their written informed consent before participation. Participation was accounted for as partial fulfillment of course requirements. For organizational reasons both groups were recruited from different courses, and were examined at different times about 6 months apart by different experimenters.

Six subjects did not complete the experiment, because they were unable to recall the criterion task under free recall conditions. These subjects were excluded from further analysis. Only the remaining 18 participants entered the final analyses reported in the following sections. As a consequence the first group (in the following labeled the Morning-Evening-Morning (MEM) group according to the experimental design; cf. “Design and Procedure” Section) comprised only 8 participants (22.1 ± 2.4 years, 4 females, one left handed, 4 males), while 10 participants (22.9 ± 1.5 years, 4 females, 6 males, one left handed) remained in the second group (labeled the Evening-Morning-Evening (EME) group accordingly).

There was no additional reward or remuneration. Participants were required to refrain from daytime naps, alcohol, excessive

caffeine-intake, and any other drugs from 24 h before initial training until the end of the experiment. Physical activity (e.g., sport practice) was permitted. All participants were naïve with respect to the criterion task and the research hypotheses.

Duration and quality of each subjects' sleep during the night of the experiment was assessed with a standardized sleep questionnaire (Goertelmeyer, 1986). There was no indication of poor sleep quality for any of the participants. Also, daytime activities during the wakening retention interval were assessed with a time-line protocol. Again, no peculiarities were observed with respect to any of the subjects.

Task and Apparatus

The criterion task required participants to repeatedly carry out a fixed sequence of 10 reaching movements with their non-dominant arm. Subjects were seated comfortably in a height-adjustable chair in front of a table-mounted electronic pegboard and a vertical computer screen with their upper trunk against the backrest. With their hand visible all the time, participants could freely move shoulder, elbow and wrist. On each trial, following a start signal they had to successively fit a small hand-held peg into the respective target-holes (depth: 22.22 mm; diameter: 12.7 mm) on the pegboard (see **Figure 1**). Thereby they followed a fixed pattern of end-point locations in the horizontal plane, which was void of any apparent regularity. Transport movements differed in amplitude (range: 3.83–33.75 cm) and direction. Precision requirements for all sequence elements amounted to an index of difficulty (ID) of 5.03 (± 0.94) on average (Fitts, 1954). According to Fitts, the ID is determined by the equation $\text{Log}_2(2A/W)$, where A represents the movement amplitude measured from one target center to the other target center and W represents the width of the target area in the direction of the movement.

The sequence to perform was never presented entirely before or during execution. Rather, participants learned the sequence by repeated execution, similarly to a serial reaction time task. During acquisition, targets were visualized one

after the other prior to each reaching movement on a computer screen. Correct execution of a sequence element was indicated by a color change of the respective target stimulus from red to green, while the next target symbol was illuminated red. In case of a reaching error, the symbol representing the target that had been missed turned green as well, while the next target was illuminated red. Thus, explicit error control always required participants to compare the peg's present position on the pegboard to the target position indicated on the screen. As soon as one sequence element was terminated, the next reaching movement had to be started immediately, until the sequence was completed. Once a sequence trial was finished, subjects had to place the peg back into the starting position and prepare themselves for the next trial. After announcing they were ready again participants received an oral start-signal about 1 s later, and then executed the next trial. This procedure was repeated until a block of 10 trials had been accomplished. During recall, no extrinsic information (neither stimulus information nor feedback) was provided. Sequence configuration, raw data assessment and screen display during sequence execution were controlled by means of LMD Software (Wagner: IAT Leipzig, Germany).

Dependant Measures

Acquisition and recall tests were organized in successive blocks of 10 trials, separated by 30-s resting periods. To prevent any build up of fatigue during acquisition, the resting period following block six was extended to 2.5 min. Performance measures were number of Erroneous Sequences (ES) per trial block (i.e., error rate), and Total Execution Time (TET) per sequence, with TET averaged for each subject across correct sequences in a trial block. TET thereby is inversely proportional to sequence execution speed. Participants were instructed to execute each single sequence-trial as rapidly as possible with as few errors as possible. They were also advised not to speed up performance at the expense of an increasing error rate. Instructions were followed by most of the participants,

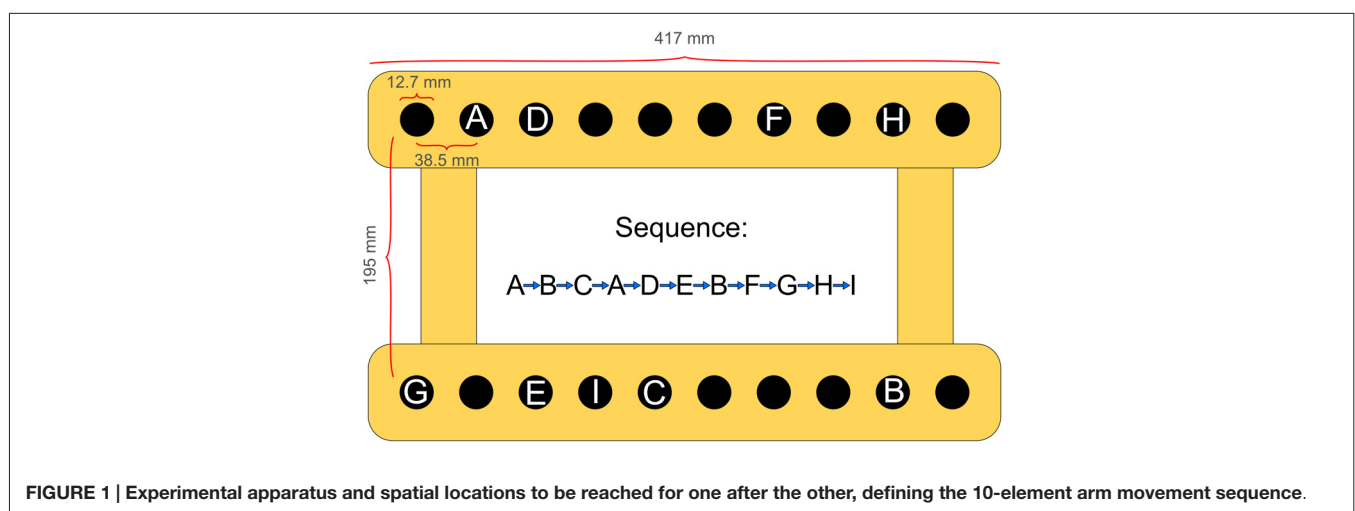


FIGURE 1 | Experimental apparatus and spatial locations to be reached for one after the other, defining the 10-element arm movement sequence.

resulting in marked skewness of the dependent variable ES (i.e., number of ES).

It should be mentioned here that this gross motor task was sufficiently complex and difficult to prevent performance reaching an asymptote within one single practice session. As had been shown previously in a pilot study with eight subjects (23.13 ± 2.1 years, 4 females, 4 males) extensively practicing this same criterion task on three successive days (600 trials altogether; two training sessions of 100 trials per day, stimulus information continuously provided), mean performance (i.e., sequence execution speed, operationalized via TET) continuously increased following a power function, and started to level off only after about 550 trials at about 5.7 s TET on average (unpublished data; Schmitz and Waßmuth, 2013). It also became clear from that study that more than 100 trials would be needed to fully memorize the spatial movement pattern.

Design and Procedure

After being shortly familiarized with the electronical pegboard and the peg-plugging procedure in general, both experimental groups received initial training of the criterion task (12 blocks of 10 trials each). Both groups then were retested three times in a free-recall condition, namely 15 min after end of practice (Post-Training), and again 12 h (Retest 1) and 24 h later (Retest 2), with each Retest comprising three blocks of 10 trials. The first group to take part in this experiment received initial training in the morning (7–9 a.m.) and was labeled the MEM group accordingly, while the second group practiced in the evening (7–9 p.m.), and was labeled the EME group respectively. Thus, subjects in the MEM-group had a regular night's sleep during their second 12-h retention interval, those in the EME-group during their first 12-h retention interval (cf. **Figure 2**). To prevent mental rehearsal of the criterion task during the 15-min retention interval directly following acquisition, participants were asked to read a series of comic stories combining pictures and text. They also were instructed to report on the stories' content at the end of the respective test session.

Statistics

Changes in performance during acquisition and retention were analyzed with reference to five different time points, namely “Start of Practice”, “End of Practice”, “Post-Training”, “Retest 1”, and “Retest 2”. Time point-specific performance values were calculated as follows: first, for each subject ES- and TET-measures were averaged across trials per block. Then for each subject and dependent variable, average performance measures were calculated from the first three initial training blocks (Start of Practice, blocks 1, 2 and 3) and from the last three initial training blocks (End of Practice, blocks 10, 11 and 12), while retest measures were calculated from blocks 13, 14, and 15 (Post-Training), 16, 17, and 18 (Retest 1), and 19, 20, and 21 (Retest 2) respectively. Group mean values (medians) were calculated on this basis.

In the presence of small sample sizes and extreme skewness of the dependent variable ES for inferential statistics non-parametric procedures were applied. Accordingly, Friedman test and Wilcoxon test were used for within-group comparisons, while Mann-Whitney *U* test was applied when data were compared across groups. A significance level of $p < 0.05$ was used for all inferential statistics. In case of multiple testing Bonferroni-corrections were applied. As a rule statistical significance was assessed two-tailed, with exact *p*-values being reported. Effect sizes were provided in terms of Cohen's *r*

$$\left(r = \frac{|z|}{\sqrt{N}} \right) \text{ and } \Phi_c \left(\Phi_c = \sqrt{\frac{\chi^2}{N(k-1)}} \right)$$

with respect to non-parametric tests (Fritz et al., 2012).

RESULTS

Descriptive Data

Performance data (i.e., number of ES and TET) achieved by each group at the respective time points are presented in **Table 1**.

Acquisition and Transfer to Free Recall

In a first step, changes in performance during acquisition and at transfer to the first free-recall test were determined for both

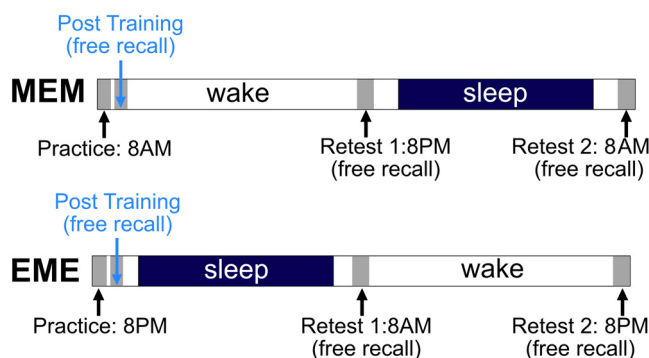


FIGURE 2 | Experimental paradigm. For details see text.

TABLE 1 | Behavioral data: number of Erroneous Sequences (ES) and Total Execution Time (TET).

Time point	MEM-Group		EME-Group	
	ES [n]	TET [s]	ES [n]	TET [s]
Start of practice	2.16 (0.7–3.5)	10.33 (9.8–11.2)	1.33 (0.3–2.9)	11.78 (11.2–12.7)
End of practice	0.33 (0.0–1.5)	7.64 (6.8–8.4)	0.83 (0.5–1.3)	9.39 (7.9–10.6)
Post training (free recall)	1.00 (0.0–1.2)	7.50 (6.7–8.6)	1.00 (0.5–2.3)	9.41 (8.4–11.3)
Retest 1 (free recall)	0.50 (0.3–2.0)	7.55 (6.4–8.5)	0.66 (0.3–1.5)	8.19 (7.0–10.5)
Retest 2 (free recall)	0.50 (0.0–1.2)	6.57 (6.1–7.5)	0.50 (0.0–1.3)	7.98 (7.1–9.8)

Reported are medians and lower and upper quartile values (in parentheses) for Number of Erroneous Sequences (ES) and Total Execution Time (TET) from each experimental group at each time point. MEM, Morning (Acquisition and Post-Training) – Evening (Retest 1)–Morning (Retest 2); EME, Evening (Acquisition and Post-Training) – Morning (Retest 1) – Evening (Retest 2). Shaded areas indicate sleep-filled retention periods.

groups. Throughout acquisition, number of *ES* was low on average ($MD = 0.87$) in the MEM-group, but even so from Start of Practice to End of Practice error rate significantly decreased ($Z = -2.527$, $p = 0.008$, Cohen's $r = 0.892$), as did *TET* ($Z = -2.521$, $p = 0.008$, Cohen's $r = 0.890$). However, when participants were subjected to the first free-recall test at Post-Training, compared to End of Practice both *ES* ($Z = -0.527$, $p = 0.688$) and *TET* ($Z = -0.280$, $p = 0.844$) statistically remained about the same. Also in the EME-group, error rate was low on average throughout acquisition ($MD = 1.16$). While number of *ES* this time did not change significantly from Start of Practice to End of Practice ($Z = -1.602$, $p = 0.129$), *TET* again significantly decreased ($Z = -2.521$, $p = 0.008$, Cohen's $r = 0.890$). When participants underwent the first free-recall test at Post-Training, compared to End of Practice both *ES* ($Z = -1.266$, $p = 0.258$) and *TET* (EME: $Z = -0.968$, $p = 0.375$) statistically remained about the same again.

Thus, both groups during acquisition significantly improved sequence execution speed and also somewhat reduced error rate, while transfer from an informational guided practice condition to free recall 15 min later did not yield any performance decrements. On the whole, error rate was real low throughout the whole experiment in either group, and there was no speed-accuracy trade-off across time points.

Retention (Free Recall Only)

In a second step possible performance changes during retention under free-recall conditions had to be determined. According to our theoretical considerations it was of specific interest, if possible performance changes during the sleep-filled retention intervals were any different from performance changes during the respective wake intervals. Considering the small sample sizes, and in order to achieve sufficient statistical power, we applied the following procedure: data of both experimental groups were combined and subjected to the respective statistical tests conjointly, so that pre- and post-wake performance data of all 18 participants could be compared directly, and pre- and post-sleep performance data of all 18 participants could be compared directly, too. Due to the circadian offset of 12 h between both experimental groups the combined pre- and post-wake interval and pre- and post-sleep interval data for each dependent variable

had to be compared in two separate test runs. It has been argued that these two tests were conceptually related. Therefore the level of significance in these cases was and set at $p = 0.025$ (two-tailed) following Bonferoni correction.

The following results now refer to the combined data of both groups. According to the respective Wilcoxon tests, *error rate* (*ES*) remained the same across both retention intervals (wake retention interval: $p = 0.404$; sleep-filled retention interval: $p = 0.106$). However, *sequence execution time* (*TET*) significantly decreased during the sleep-filled retention interval ($Z = -3.245$, $p = 0.001$, Cohen's $r = 0.540$), but not so during the wake retention interval ($Z = -1.894$, $p = 0.060$, Cohen's $r = 0.315$). The respective *TET*-data are depicted in **Figure 3**.

Thus, regarding our total sample ($N = 18$) the following became evident: *TET* significantly decreased (i.e., sequence execution speed increased) during the sleep-filled 12-h retention interval, but remained statistically unchanged during the respective 12-h wake interval. *Error rate* (*ES*), on the other hand, remained completely unaffected by the respective treatment conditions throughout retention. So also during retention there was no indication of any speed-accuracy trade-off. These findings were well in line with our theoretical expectations of sleep-dependent offline-gains in sequence execution speed. They were also corroborated by statistical analysis at the single group level (see “Supplementary Material”).

DISCUSSION

The present study was intended to test the notion of sleep-related offline learning coming into effect in a sequentially structured gross motor task after only limited practice and under free recall requirements. These are conditions common to many applied areas in the motor learning domain. From a theoretical point of view, any offline improvements in performance observed at retention under these conditions can be attributed solely to an enhanced sequence memory, since continued online learning at retests is effectively prevented by the absence of criterion-related stimulus information. In traditional motor learning experiments, only terminal feedback is usually removed to prevent further learning. But as long as stimulus information is still present at retesting (like e.g., in the typical serial reaction time task), continued updating

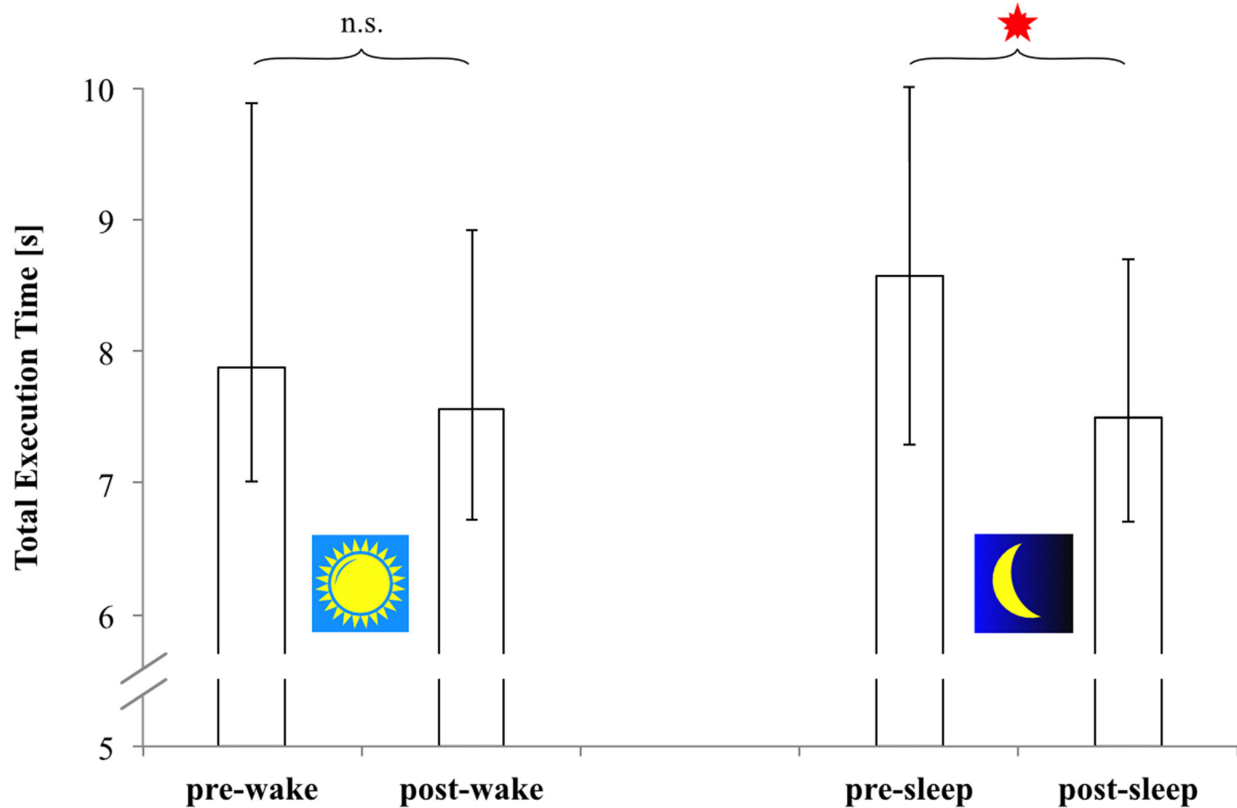


FIGURE 3 | Total execution time (seconds; correct sequences only) of all 18 subjects (Morning-Evening-Morning-group and Evening-Morning-Evening-group combined) at free recall. Presented are measures based on the combined data from both groups' pre- and post-wake retention tests (left panel), and from both groups' pre- and post-sleep retention tests (right panel). Open bars: medians; Error bars: upper and lower quartiles. *Significant difference of group medians ($p = 0.001$).

of sequence memory on grounds of externally provided information cannot be prevented. From an ecological point of view such testing conditions are not likely to reliably engage retrieval strategies relevant to many real-life situations in the field.

In the present study a 10-element sequence of reaching movements was used for a criterion task. Participants executed this sequence on an electronic pegboard with their unrestrained non-dominant arm, thereby following a fixed spatial pattern in the horizontal plain. The pattern had no apparent regularities. The sequence had to be carried out as rapidly and with as few errors as possible. Dependent variables were number of ES, and total sequence execution time. These performance measures thus represented error rate and sequence execution speed. Two groups of altogether 24 subjects initially learned this sequence for a total of 120 trials either in the morning (MEM-group) or in the evening (EME-group). Performance was retested 15 min post training, as well as 12 h and 24 h later. Thus, each group was provided with one sleep-filled and one wake retention interval. All three retests required free recall of the criterion sequence.

At the end of practice all subjects had more or less explicit knowledge of the sequence they had learned, and were using different retrieval strategies at (re)testing. This can be concluded

from subjects' verbal reports given at the end of the experiment. However, to which extent participants used cognitive retrieval strategies or more procedural aspects of the motor task in question (cf. Hikosaka et al., 1999) cannot be decided. At any case, six subjects (four in the MEM- and two in the EME-group) were unable to reproduce the initially learned sequence under free recall conditions, even when they tried to explicitly remember the sequence. These subjects were excluded from further analysis.

In the remaining 18 subjects error rate was low right from the beginning and dropped to well below one erroneous sequence per block of 10 trials at the end of practice. Sequence execution speed improved significantly in both groups during acquisition. During retention error rate did not change any more (no group differences). Total sequence execution time during retention significantly *decreased* following sleep, but not following wake. This held true for the total sample, and could also be corroborated for each group separately (cf. "Supplementary Material"). Throughout the experiment there was no speed-accuracy trade-off.

It should be noted that sequence execution time at the end of practice in both groups was still well above (at least 2 s) asymptotic performance level. The performance asymptote for

this same task has been determined in a previous study after three days of continued practice by eight young subjects of comparable age (Schmitz and Waßmuth, 2013). Therefore it seems unlikely that global differences in sequence execution speed between experimental groups could have biased the sleep-related improvements in performance found at retention to any relevant extent. Also, this finding of sleep-related motor performance improvement was independent from retention interval duration and time of day of learning: the EME-group initially acquired the criterion sequence in the evening and was afforded sleep during the first 12 h retention period. The MEM-group to the contrary learned the sequence in the morning and slept during the second 12 h retention period. All in all these results corroborate recent findings of sleep-related motor offline learning in a very similar task, however with the same stimulus information provided at retention as well as during the initial learning phase (Malangré et al., 2014).

It should be mentioned that in the EME-group, following significant sleep-dependent offline improvement, sequence execution time also decreased somewhat during the second (the wake) retention interval. This effect is close to significance ($p = 0.064$, Cohen's $r = 0.597$; see "Supplementary Material"), and was not observed in the MEM-group. From this one might conjecture that sleep-dependent consolidation mechanisms are still in process during the following waking period, while this is not the case during the waking period prior to sleep. This aspect certainly requires closer consideration in the future.

In this context, also the following observation might be of particular interest: in a pilot study (unpublished data) we conducted in our laboratory preceding the experiment presented in this article, two randomized groups of participants (all students at the department of sport science) practiced the same criterion task as was used in our present study either in the morning (ME-group; 21.0 ± 2.4 years; 5 females; 4 males) or in the evening (EM-group; 21.0 ± 0.98 years; 4 females; 7 males) for 120 trials, and were retested under free recall conditions 12 h later, i.e., on the same evening or on the next morning respectively. Note that there was no early free recall test shortly following acquisition. During acquisition total sequence execution time significantly decreased in either group from 9.82 s on average to 7.53 s on average. But then in this pilot study at free recall seven out of the nine subjects in the ME-group were unable to reproduce the criterion sequence after their 12 h waking interval. Obviously during a 12 h wake retention interval they had forgotten essential sequence components (i.e., certain elements and/or order of elements). To the contrary only two out of the eleven subjects in the EM-group failed to recall the sequence after their 12-h sleep-filled retention interval. Thus, sleep appeared to prevent sequence

memory to deteriorate. Also, and different from our present results, in the absence of an early free-recall test in the remaining nine subjects of the pilot-study's EM-group sequence execution speed at free recall following a night of sleep appeared to be stabilized, but not improved as compared to performance at the end of acquisition. Thus, it could be argued that withdrawing stimulus information and feedback opportunity during testing might have hidden possible sleep-dependent performance improvements.

Thus, implementation of an early free recall test (Post-Training) in our present experiment not only provided for an appropriate datum point subjects' performance at the two later free recall tests could be related to i.e., transfer-appropriate processing; cf. Lee (1988). We conjecture that it also served as a means to effectively reduce the tendency for sequence representation to decay over a 12 h waking period, and to provide a basis for subsequent enhancement of sequence memory during sleep. We assume that the necessity of free recall soon after acquisition stabilizes and even considerably elaborates the multifaceted sequence representation still intact at that point of time. This positive effect of early retesting on long term retention has recently been found for verbal material (Roedinger and Karpicke, 2006) as well as for effector transfer in motor sequence learning, which is indicative for the generalization of the abstract spatial sequence pattern (Boutin et al., 2013). Thus, testing conditions not only boosts memory when learners are allowed to practice between testing sessions as in the study of Boutin et al. (2013), but early testing under free recall conditions might also shape sequence memory so to enhance later retention.

All in all, while with the present experiment we successfully corroborated and extended recent findings on sleep-related offline learning in gross motor sequence learning tasks, there are also clear limitations to our study in that sample size was rather small, and subjects were not randomly assigned to the experimental groups.

AUTHOR CONTRIBUTIONS

AM and KB contributed extensively and equally to the work presented in this manuscript, developed the research topic and designed the experiment, prepared, analyzed data and discussed results, wrote the article and discussed and commented on the manuscript at all stages. AM supervised data collection.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fnhum.2016.00134/abstract>

REFERENCES

- Albouy, G., Fogel, S., King, B. R., Laventure, S., Benali, H., Karni, A., et al. (2015). Maintaining vs. enhancing motor sequence memories: respective roles of striatal and hippocampal systems. *Neuroimage* 108, 423–434. doi: 10.1016/j.neuroimage.2014.12.049
- Al-Sharman, A., and Siengsukon, C. F. (2013). Sleep enhances learning of a functional motor task in young adults. *Phys. Ther.* 93, 1625–1635. doi: 10.2522/ptj.20120502
- Blischke, K., Erlacher, D., Kresin, H., Brueckner, S., and Malangré, A. (2008). Benefits of sleep in motor learning - prospects and limitations. *J. Hum. Kinet.* 20, 23–35. doi: 10.2478/v10078-008-0015-9

- Born, J., and Wilhelm, I. (2012). System consolidation of memory during sleep. *Psychol. Res.* 76, 192–203. doi: 10.1007/s00426-011-0335-6
- Boutin, A., Panzer, S., and Blandin, Y. (2013). Retrieval practice in motor learning. *Hum. Mov. Sci.* 32, 1201–1213. doi: 10.1016/j.humov.2012.10.002
- Doyon, J., Korman, M., Morin, A., Dosti, V., Hadj, T. A., Benali, H., et al. (2009). Contribution of night and day sleep vs. simple passage of time to the consolidation of motor sequence and visuomotor adaptation learning. *Exp. Brain Res.* 195, 15–26. doi: 10.1007/s00221-009-1748-y
- Fischer, S., Nitschke, M. F., Melchert, U. H., Erdmann, C., and Born, J. (2005). Motor memory consolidation in sleep shapes more effective neural representations. *J. Neurosci.* 25, 11248–11255. doi: 10.1523/jneurosci.1743-05.2005
- Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *J. Exp. Psychol.* 47, 381–391. doi: 10.1037/h0055392
- Fritz, C. O., Morris, P. E., and Richler, J. J. (2012). Effect size estimates: current use, calculations and interpretation. *J. Exp. Psychol. Gen.* 141, 2–18. doi: 10.1037/a0024338
- Genzel, L., Quack, A., Jäger, E., Konrad, B., Steiger, A., and Dresler, M. (2012). Complex motor sequence skills profit from sleep. *Neuropsychobiology* 66, 237–243. doi: 10.1159/000341878
- Goertelmeyer, R. (1986). “Schlaf-Fragebogen A und B (Sf-A, Sf-B),” in *Internationale Skalen Fuer Psychiatrie* ed. C. I. P., Sclarum 4, Weinheim: Beltz.
- Gudberg, C., Wulff, K., and Johansen-Berg, H. (2015). Sleep-related motor memory consolidation in older adults depends on task demands. *Neurobiol. Aging* 36, 1409–1416. doi: 10.1016/j.neurobiolaging.2014.12.014
- Hikosaka, O., Nakahara, H., Rand, M. K., Sakai, K., Lu, X., Nakamura, K., et al. (1999). Parallel neural networks for learning sequential procedures. *Trends Neurosci.* 22, 464–471. doi: 10.1016/s0166-2236(99)01439-3
- Hoedlmoser, K., Birklbauer, J., Schabus, M., Eibenberger, P., Rigler, S., and Mueller, E. (2015). The impact of diurnal sleep on the consolidation of a complex gross motor adaptation task. *J. Sleep Res.* 24, 100–109. doi: 10.1111/jsr.12207
- Huber, R., Ghilardi, M. F., Massimini, M., and Tononi, G. (2004). Local sleep and learning. *Nature* 430, 78–81. doi: 10.1038/nature02663
- Kempler, L., and Richmond, J. L. (2012). Effect of sleep on gross motor memory. *Memory* 20, 907–914. doi: 10.1080/09658211.2012.711837
- Lee, T. D. (1988). “Transfer-appropriate processing: a framework for conceptualizing practice effects in motor learning,” in *Complex Movement Behaviour: “The” Motor-Action Controversy*, Vol. 50, eds O. G. Meijer and K. Roth (Amsterdam: Elsevier Science Ltd.), 201–215.
- Malangré, A., Leinen, P., and Blischke, K. (2014). Sleep-related offline learning in a complex arm movement sequence. *J. Hum. Kinet.* 40, 7–20. doi: 10.2478/hukin-2014-0002
- Morita, Y., Ogawa, K., and Uchida, S. (2012). The effect of a daytime 2-hour nap on complex motor skill learning. *Sleep Biol. Rhythms* 10, 302–309. doi: 10.1111/j.1479-8425.2012.00576.x
- Robertson, E. M., Pascual-Leone, A., and Press, D. Z. (2004). Awareness modifies the skill-learning benefits of sleep. *Curr. Biol.* 14, 208–212. doi: 10.1016/j.cub.2004.01.027
- Roedinger, H. L., III and Karpicke, J. D. (2006). Test-enhanced learning. *Psychological Science* 17, 249–255. doi: 10.1007/springerreference_302024
- Schmitz, L., and Waßmuth, N. (2013). *Zum Einfluß der Übungsrate auf Ausführungsgeschwindigkeit und Reproduktionssicherheit einer komplexen Armbewegungsfolge - Eine Laborstudie*. Unpublished Bachelors’ Thesis; Saarland University, Department of Sport Science.
- Walker, M. P. (2005). A refined model of sleep and the time course of memory formation. *Behav. Brain Sci.* 28, 51–104. doi: 10.1017/s0140525x05000026

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Malangré and Blischke. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution and reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Task Complexity Modulates Sleep-Related Offline Learning in Sequential Motor Skills

Klaus Blischke* and Andreas Malangré

Laboratory of Training Science, Department of Sport Science, Training Science, Saarland University, Saarbrücken, Germany

Recently, a number of authors have advocated the introduction of gross motor tasks into research on sleep-related motor offline learning. Such tasks are often designed to be more complex than traditional key-pressing tasks. However, until now, little effort has been undertaken to scrutinize the role of task complexity in any systematic way. Therefore, the effect of task complexity on the consolidation of gross motor sequence memory was examined by our group in a series of three experiments. Criterion tasks always required participants to produce unrestrained arm movement sequences by successively fitting a small peg into target holes on a pegboard. The sequences always followed a certain spatial pattern in the horizontal plane. The targets were visualized prior to each transport movement on a computer screen. The tasks differed with respect to sequence length and structural complexity. In each experiment, half of the participants initially learned the task in the morning and were retested 12 h later following a wake retention interval. The other half of the subjects underwent practice in the evening and was retested 12 h later following a night of sleep. The dependent variables were the error rate and total sequence execution time (inverse to the sequence execution speed). Performance generally improved during acquisition. The error rate was always low and remained stable during retention. The sequence execution time significantly decreased again following sleep but not after waking when the sequence length was long and structural complexity was high. However, sleep-related offline improvements were absent when the sequence length was short or when subjects performed a highly regular movement pattern. It is assumed that the occurrence of sleep-related offline performance improvements in sequential motor tasks is associated with a sufficient amount of motor task complexity.

Keywords: sleep, memory consolidation, offline learning, task complexity, gross motor task, motor sequence learning

OPEN ACCESS

Edited by:

Juliana Yordanova,
Institute of Neurobiology
(BAS), Bulgaria

Reviewed by:

James Ward Antony,
Princeton University, United States
Robert Stickgold,
Harvard Medical School,
United States

*Correspondence:

Klaus Blischke
k.blischke@mx.uni-saarland.de

Received: 05 December 2016

Accepted: 03 July 2017

Published: 25 July 2017

Citation:

Blischke K and Malangré A
(2017) Task Complexity Modulates
Sleep-Related Offline Learning in
Sequential Motor Skills.
Front. Hum. Neurosci. 11:374.
doi: 10.3389/fnhum.2017.00374

INTRODUCTION

For two decades, there has been mounting evidence showing that sleep plays a crucial role in learning and memory consolidation (Smith, 1995; Stickgold et al., 2001; Walker, 2005; Genzel et al., 2014). Drawing on the concept of an “active system consolidation” process (Born and Wilhelm, 2012), it is assumed that newly encoded skill representations are being actively (and repeatedly) reprocessed during sleep, the resulting in a qualitative reorganization and long-term stabilization of the respective memory representations.

In this context, sleep has also been implicated in the consolidation process of *motor* skill memory following initial acquisition, with delayed learning being achieved in the absence of further practice. Specifically regarding motor *sequence* learning, significant offline improvements in speed and accuracy have repeatedly been found following nocturnal and diurnal sleep, whereas no such benefits were provided by equivalent wake periods (Walker et al., 2002; Fischer et al., 2005; Doyon et al., 2009; Malangré et al., 2014).

It should be noted, however, that the sleep-based enhancement hypothesis has recently been called into question. In several studies, researchers have identified a number of moderating variables that may account for at least a portion of the performance improvement following sleep (see Pan and Rickard, 2015, for a comprehensive review and meta-analysis). Some of these data even suggest that under some circumstances, skill performance after sleep may not be better than after a period of wakefulness (e.g., Backhaus et al., 2016), inviting a reconsideration of sleep's theoretical role in the consolidation of procedural memories (Nettersheim et al., 2015). This debate is still unresolved at present.

As has been found repeatedly in motor sequence learning studies, sleep specifically enhances an allocentric spatial sequence representation, whereas an egocentric “motor” representation progressively develops with time (Cohen et al., 2005; Witt et al., 2010; Albouy et al., 2013). With regard to the neural substrates underlying the formation of these two components of sequence memory, recent fMRI-studies have shown that the abstract spatial map representing the schematic “gist” of motor sequence memory, is supposedly created through hippocampo-cortical activity during initial acquisition, while the motoric representation is assumed to be supported by the striato-cortical system. Antagonistic dynamic activity between the two systems during initial motor learning then is thought to condition subsequent sleep-related processes underlying enhanced sequence memory consolidation (Albouy et al., 2015).

Thus, it appears that sleep predominantly affords facilitation of the abstract, cognitive memory trace of a sequence, driven by early activity in the hippocampus and the frontal and parietal cortices. Development of this memory component is usually accompanied by at least rudimentary awareness of some of the regular features inherent to the sequence being learned. Therefore, this memory component is supposedly associated with declarative knowledge concerning the action's goal as well as the type of sequence elements and their temporal order. Most likely, further declarative knowledge aspects of the task are liable to become part of this memory representation. This notion is also supported by the observation that subsequent sleep-dependent enhancement in performance has been consistently reported in *explicit* sequence learning conditions. In the respective paradigms, the sequence of elements to be performed is explicitly provided to the participants either prior to or throughout the initial practice (Albouy et al., 2013; Malangré et al., 2014). To the contrary, however, sleep appears to play no such critical role in implicit motor sequence learning (Robertson et al., 2004; Song et al., 2007; Nemeth et al., 2010).

Regarding explicitly acquired motor sequences, converging evidence from the behavioral sciences and neuroscience has uncovered some basic principles of sleep-dependent memory consolidation and continues to provide a deeper understanding of the underlying processes and mechanisms. At the same time, however, to date, surprisingly little attention has been paid to the question as to what extent these findings can be generalized across the domain of sequential motor tasks. Only recently have a number of studies deliberately incorporated *gross* motor tasks (Genzel et al., 2012; Kempler and Richmond, 2012; Morita et al., 2012; Al-Sharman and Siengsukon, 2013; Malangré et al., 2014; Gudberg et al., 2015; Hoedlmoser et al., 2015; Malangré and Blischke, 2016). Here, it has repeatedly been argued that nearly all the paradigms used previously to study sleep-dependent offline motor learning have been limited to relatively “simple” tasks (e.g., finger-to-thumb opposition tasks; serial reaction time tasks; sequential finger tapping tasks) compared to the motor tasks usually performed in daily life and thus might have only limited implications for applied areas such as vocational training, sports and rehabilitation. To overcome this limitation, in these recent studies, gross motor tasks involving unrestrained whole limb movements were introduced, including the upper and lower extremities.

Down the line, this choice of tasks was motivated by the respective authors' objective to sufficiently increase the *task complexity* to adequately meet the functional requirements of real world applications. In this context, however, the concept of task complexity was usually conceived in a rather general way. For the most part and only vaguely defined, it has been associated with such diverse aspects such as, for example, sequence length, kinematic constraints, redundancy control, inter-limb coordination, the need to plan and execute goal-directed movements in Euclidian space, and reactions to different environmental (i.e., visual, auditory and proprioceptive) stimuli. However, so far as the extent to which these various aspects genuinely exercise influence over sleep-dependent motor memory consolidation processes remains open. One might even ask whether the notion of an all-encompassing concept of task complexity proves helpful at all with regard to elucidating the link between task demands and sleep-dependent memory consolidation in any sufficient detail.

This being said, surprisingly little effort has been undertaken to systematically examine the role of task complexity in sleep-dependent offline motor learning. To our knowledge, task complexity has been manipulated experimentally in only one study so far (Kuriyama et al., 2004). In this study, four groups of participants initially practiced uni- and bimanual key-pressing sequences of different lengths (five respective nine elements) and were retested 24 h later after a night of sleep. To produce the five-element configuration, the subjects had to press four numeric keys either with their left hand (unimanual group) or with two fingers from each hand (bimanual group). For the nine-element configuration, one group of subjects again pressed the four numeric keys with their left hands, while the corresponding bimanual group had to press eight numeric keys using four fingers from each hand. All groups significantly improved their sequence execution speed during initial training

and then went on to significantly improve overnight. While the amount of offline improvement was similar for both of the short five-element sequences independently of number of hands involved, increasing the sequence length to nine elements in the unimanual condition resulted only in modest, but not significantly greater, offline improvements, whereas performing the nine-element bimanual task that used all eight digits produced dramatically greater overnight increases in execution speed compared to all the other groups. The authors attributed this significant result to the combined demand of memory load (i.e., sequence length) and the extent of movement coordination (i.e., the number of hands and digits required) and thus concluded that the more “complex” a task is, the larger the degree of sleep-dependent offline learning is.

Taking these propositions as a starting point to further elucidate the relationship between task complexity and motor memory consolidation, in our present work, we tried to (a) disentangle different complexity components, which are normally intertwined in many motor acts, and we also (b) adopted a gross motor skill for a criterion task that incorporated functional requirements typical of numerous activities of daily living. To this end we conducted a series of three experiments. In each experiment, the participants were to learn an unrestrained arm movement sequence of different levels of complexity. The task complexity was varied either with respect to *sequence length* (i.e., number of task elements) or with respect to *structural complexity* (defined by the amount of sequence regularity). Per definition, structural complexity increases as the regularity of the respective sequential pattern decreases. In the first experiment, task complexity in terms of sequence length and structural complexity was set at a level that was supposedly high enough to induce sleep-related performance enhancement. In the second experiment, sequence length was reduced by 50%. In the third experiment, structural complexity was reduced by organizing the sequence in a much more regular fashion, while sequence length was the same as in Experiment 1. Reducing the task complexity in Experiments 2 and 3 was expected to also reduce the memory load associated with learning these tasks to such an extent that sleep-related offline learning would come into effect only to a lesser degree than in Experiment 1, or maybe even not at all.

MATERIALS AND METHODS

All three experiments followed the same experimental paradigm. They were conducted in accordance with the ethical standards of the 1975 Declaration of Helsinki and were approved by the Ethics Committee of Faculty 5 Empirical Social Sciences of Saarland University.

Participants

A total of 73 young and healthy men and women participated in our studies. Of these, 24 subjects (22.77 ± 2.45 years; 12 men, one left-handed; 12 women) participated in Experiment 1. Another 24 subjects (22.38 ± 2.08 years; 13 men, two left-handed; 11 women, one left-handed) participated in Experiment 2. The remaining 25 subjects (21.64 ± 0.96 years; 16 men,

one left-handed; 9 women, two left-handed) participated in Experiment 3. In each experiment, the participants of either sex were randomly assigned to two experimental groups of 12 subjects each in Experiments 1 and 2 and groups of 12 (four female) or 13 (five female) subjects in Experiment 3. Experimental groups were always balanced as far as possible with respect to the participants' sex.

The subjects had no prior knowledge of the criterion task and were naïve about the hypotheses of the experiments. They were required to refrain from daytime naps, alcohol, excessive caffeine intake and any other drugs from 24 h before their first training session until the end of the last test. Physical activity (e.g., sport practice) was permitted. Their participation was credited as partial fulfillment of the course requirements. There was no additional reward or remuneration. Written informed consent was obtained from all participants prior to their experiment.

Tasks and Apparatus

In each of the experiments, during the task execution, the subjects were seated in front of a table-mounted *electronic pegboard* that was placed horizontally in front of them with their upper trunks strapped to the backrest. Thus, while their body position was fixated with respect to the pegboard, the participants could use their entire arms by freely moving the shoulder, elbow and wrist. The pegboard consisted of two horizontal wooden bars (41.7 cm long, 16 cm apart), each containing 10 holes that were 22.22 mm in depth, 12.7 mm in diameter, 25.4 mm apart in the left-right and 195 mm apart in the forward-backward dimensions (see Figure 1).

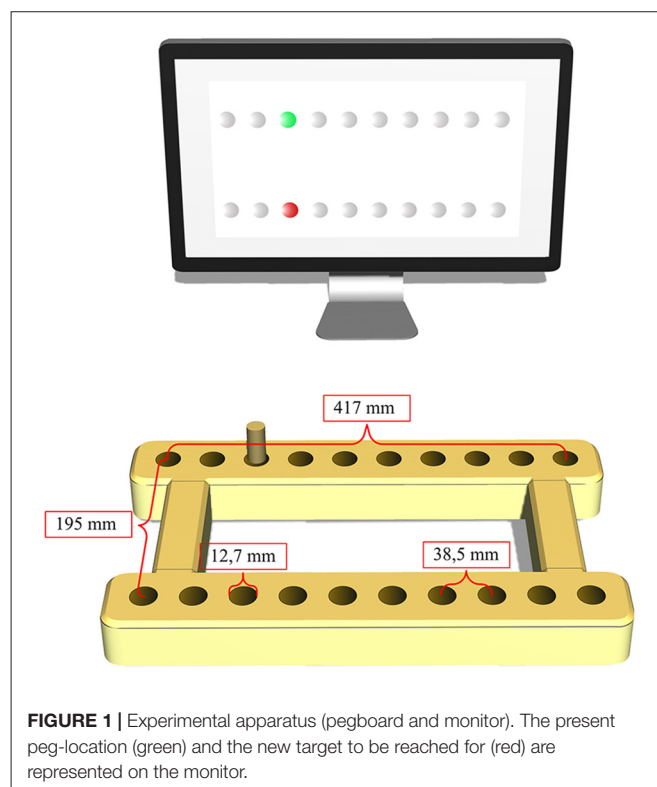
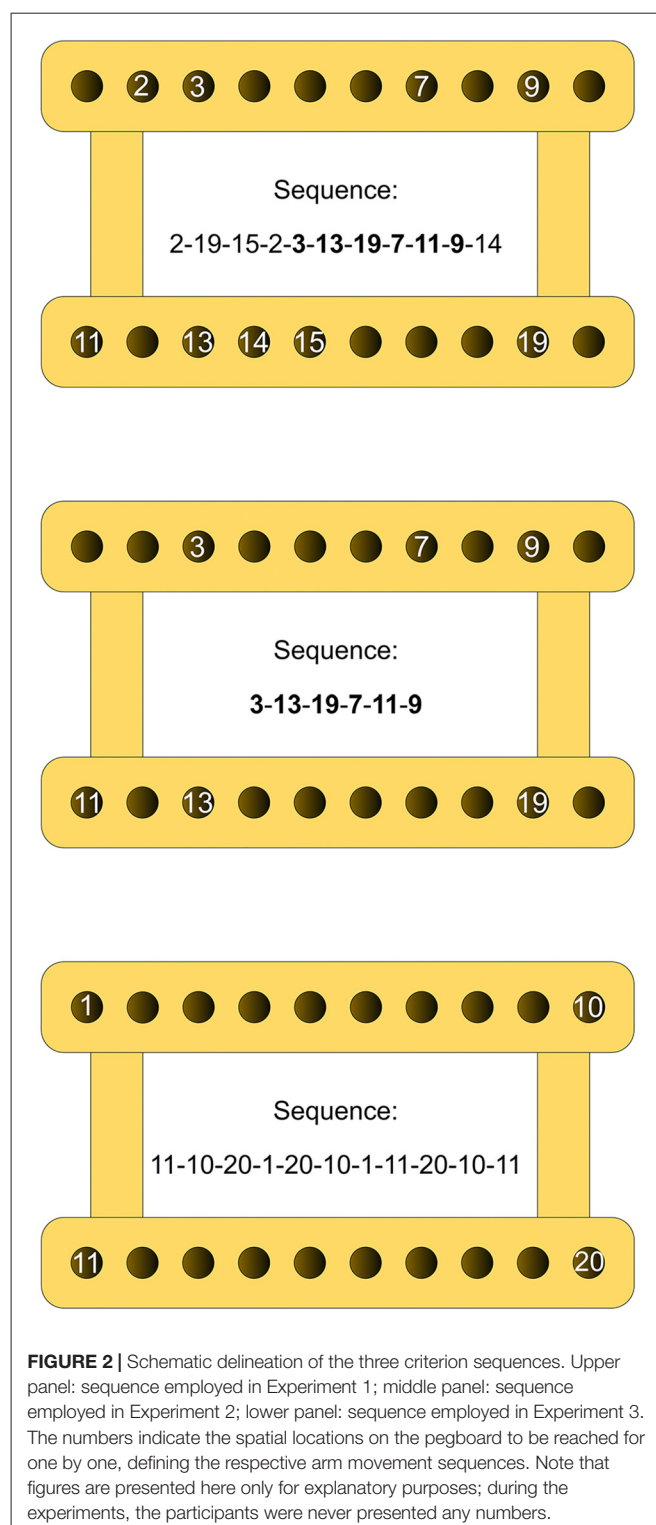


FIGURE 1 | Experimental apparatus (pegboard and monitor). The present peg-location (green) and the new target to be reached for (red) are represented on the monitor.

The *criterion tasks* employed in each of the experiments required the participants to carry out a spatially defined *arm-movement sequence*. To this end, the subjects were to execute a series of goal-directed movements with a small wooden peg (9.52 mm in diameter and 50.8 mm long) held with a pincer grip in the non-dominant hand. At the end of each reaching movement, they had to quickly fit the peg into a designated hole on the pegboard, thereby closing the magnetic contact. Once the contact was closed, the respective sequence element was terminated, and the next transportation movement had to be started immediately until the sequence was completed. Precision requirements of the reaching movements were determined by calculating their index of difficulty (ID; Fitts, 1954). This was done according to the equation $\text{Log}_2(2A/W)$, where A represents the movement amplitude measured from one target center to the other, and W represents the target width. IDs >4.5 are regarded as high (Boyle and Shea, 2011).

We used *three different criterion sequences*, one for each experiment (see **Figure 2**): (a) in *Experiment 1*, the subjects had to carry out a *ten-element* sequence, which followed a spatial pattern *void of any apparent regularity* (**Figure 2**, upper panel). Reaching movements differed with respect to direction (left, right, forward, backward and diagonal) as well as amplitude (range: 3.83 cm to 36.87 cm). No two movements were identical. Their mean ID amounted to $5.04 (\pm 0.95)$; (b) in *Experiment 2*, the subjects had to perform a *five-element* sequence (**Figure 2**, middle panel). The respective movements were identical to elements 5 through 9 of the sequence employed in Experiment 1. They also followed a spatial pattern *without any apparent regularity* and differed in direction and amplitude (range: 20.50 cm to 36.87 cm; mean ID: $5.35 (\pm 0.36)$); and (c) in *Experiment 3*, the criterion sequence again included *ten elements* (**Figure 2**, lower panel). Nine out of 10 reaching movements again differed in direction and amplitude (range: 20.50 cm to 40.10 cm). Elements five and nine were identical. This time, the mean ID was $5.57 (\pm 0.49)$ and the spatial pattern to be followed included two familiar geometric forms (an x-shape and a rectangle), which made the sequential pattern fairly regular. Additionally, to reduce ambiguity in target identification, each reaching movement terminated at one of the four pegboard corners and nowhere in between.

The participants were explicitly told that a sequence with a fixed number of reaching movements was used consistently throughout the experiment. However, the sequence was never presented entirely before or during execution. Rather, the participants had to learn the task by repeated execution. That is, during acquisition and retention, the targets were visualized one after the other prior to each reaching movement on a computer screen. Correct execution of a sequence element was indicated by a color change of the respective target stimulus from red to green, while the next target symbol was illuminated in red. In case of an error, the symbol representing the target that had been missed turned green as well, while the next target was red. Thus, explicit error control always required participants to compare the peg's present position on the pegboard to the target position indicated on the screen. Throughout one sequence trial, only the target stimulus of the one element that had just been executed was illuminated in green. Thus,



the participants were always presented with stimuli specifying the starting point (green) and end point (red) of just one sequence element at the time. When one sequence was finished, the subjects were to place the peg back into the starting position again (the first position of the sequence indicated by a green symbol on the computer screen) and mentally prepare

for the next trial's execution. After signaling that they were ready, the experimenter gave a verbal start signal about 2 s later. The next trial was then performed until a block of 10 trials had been completed. Sequence configuration, raw data assessment and the screen display during the sequence execution were controlled by means of LMD software (Wagner; IAT Leipzig).

Design and Procedure

Before initial training, in all three experiments, the subjects were briefly familiarized with the apparatus and the peg-plugging procedure in general. To this end, with their non-dominant hand, they conducted 10 trials of a three-element sequence with elements different from those of the criterion task. Following familiarization, the participants were informed about the length of the respective criterion sequence, and they were told that the sequence was invariant across trials and trial blocks. Thereafter, the participants in each experimental group received initial training in the criterion task (10 blocks of 10 trials each). To prevent fatigue, the trial blocks were always separated by 30-s intervals. The subjects were retested following a 12-h retention interval, with the retest comprising three blocks of 10 trials each. During the retest, the same stimulus information was available to the subjects as during the initial practice. In each experiment, one group of subjects received initial training in the morning (7–9 a.m.) and was retested again in the evening. This group was labeled the ME (= Morning-Evening) group. The other group practiced in the evening (7–9 p.m.), was retested the next morning, and was labeled the EM (= Evening-Morning) group correspondingly. Thus, the subjects in the ME groups had to stay awake during a 12-h retention period until the retest (no naps allowed), while those in the EM groups had a regular night's sleep during their 12-h retention interval. For subjects assigned to the EM groups, the duration and quality of each subject's sleep during the experimental night was assessed with a standardized sleep questionnaire (Goertelmeyer, 1986). For subjects assigned to the ME groups, daytime activities were assessed with a time-line protocol, which was checked at Retest. The participants were always instructed to perform each movement sequence as quickly and with as few errors as possible. However, they were also advised not to speed up performance at the expense of an increasing error rate. *Performance measures* were the number of Erroneous Sequences (ES) per trial block and Total Execution Time (TET) per sequence, with TET averaged for each subject across the correct sequences in a trial block. The TET is thereby inversely proportional to the sequence execution speed.

Data Analysis and Statistics

In all three experiments, changes in performance during the acquisition and retention were analyzed with reference to three different time points, comprising three trial blocks each. These were “Start-of-Practice” (blocks 1, 2 and 3), “End-of-Practice” (blocks 8, 9 and 10) and “Retest” (blocks 11, 12 and 13). For each individual and dependent variable, the performance measures were averaged across the respective trial blocks comprised at each time point. The group means were then calculated from the individual subjects' mean ES and

TET per time point during the initial training and retest, respectively. Moreover, to render offline performance changes from different experiments directly comparable, the change scores were calculated for the two retention periods (sleep/wake). To this end, for each participant, the difference in the TET values from “End-of-Practice” to “Retest”, divided by “End-of-Practice”, was calculated. Multiplication by 100 yielded the respective %-values. Normalization of the difference in TET performance across the Time Points (“End-of-Practice”, “Retest”) to the performance achieved by End-of-Practice also adjusts the offline performance differences to the theoretical “room for change” that may differ across individuals.

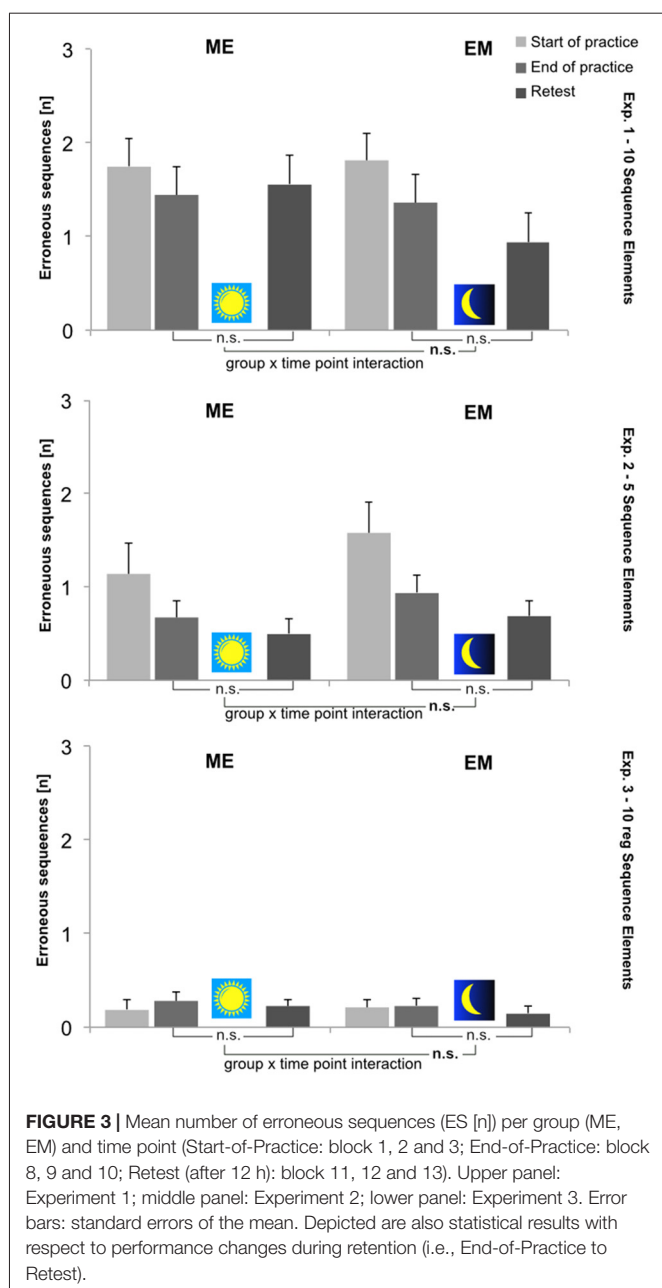
For inferential statistics, two-way analysis of variance (ANOVA) and paired *t*-tests were run. With respect to repeated-measures factors, in the case of violation of the sphericity assumption, a *df*-correction according to Greenhouse-Geisser was applied. A significance level of $p < 0.05$ was used for all inferential statistics. The effect sizes were provided in terms of η_p^2 with respect to the ANOVAs and Cohen's *d* with respect to the *t*-tests. In the case of multiple hypothesis testing using *t*-tests, the *p*-values were Bonferroni-corrected. One-sample *t*-tests were used to compare the change scores against zero. All calculations were conducted with SPSS-PC, version 15.0 (SPSS Inc., Chicago, IL, USA).

RESULTS

The average sleep duration in the EM groups amounted to 7.00 (± 0.51) h in Experiment 1, 6.50 (± 0.50) h in Experiment 2 and 6.45 (± 0.40) h in Experiment 3. There was no indication of poor sleep quality for any of the participants in any of the experiments. Additionally, no naps or peculiarities were reported for any of the participants assigned to the respective ME groups.

In all three experiments, the mean *error score* (ES) in general was low throughout acquisition (Experiment 1: 1.6 (± 0.2); Experiment 2: 0.9 (± 0.4); and Experiment 3: 0.2 (± 0.04)) and remained unchanged throughout the retention intervals. In this respect, experimental groups did not differ, nor was there a Group \times Time Point interaction in any of the studies. In all three studies, the error score remained unaffected by the treatment conditions (i.e., wake or sleep during retention). The ES results for all three experiments at each time point are shown in **Figure 3**.

Sequence execution time (TET), to the contrary, significantly decreased during acquisition in each experiment. During retention, the TET significantly decreased once more in Experiments 1 and 3. This was not the case in Experiment 2. A significant Group \times Time Point interaction for TET changes from the End-of-Practice to Retest was found only in Experiment 1, indicating a significantly greater reduction in TET at Retest in the EM group following sleep than in the ME group after waking. The TET results for all three experiments at each time point are depicted in **Figure 4**. For a more detailed inspection, for each experiment the time course of TET change across all 10 practice blocks and the three Retest blocks is shown in the Supplementary Figure S1. There

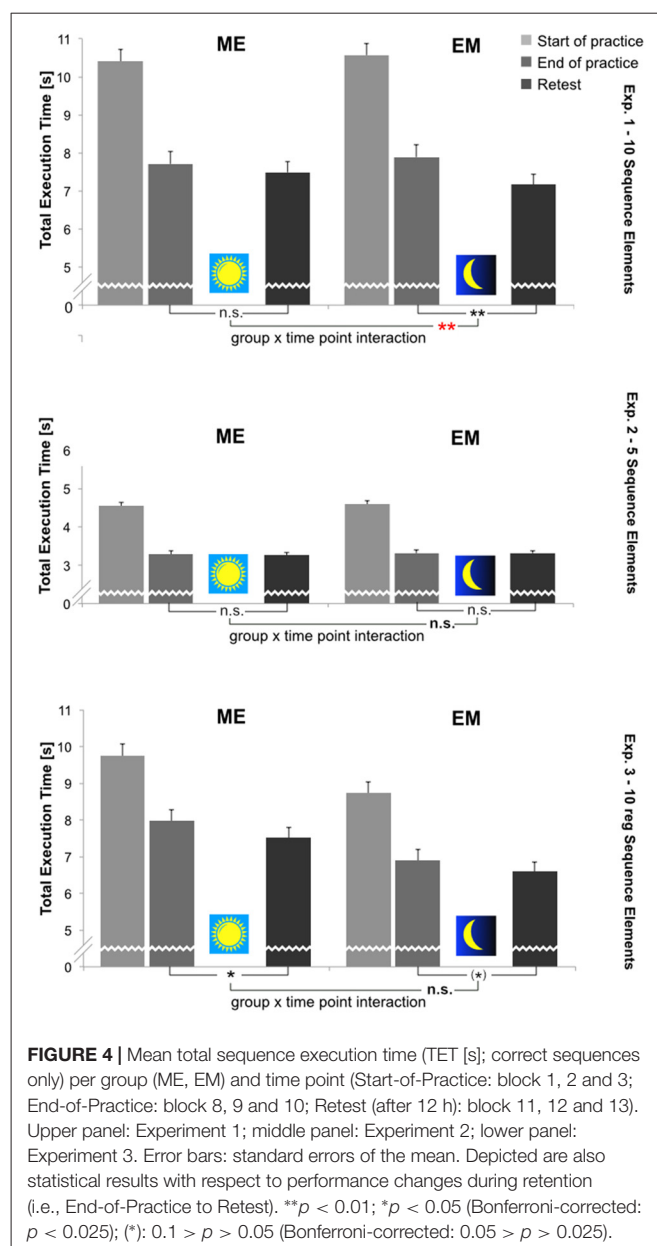


was no indication of a speed-accuracy trade-off in any of the experiments. Below, detailed statistical results are reported separately for each experiment.

Experiment 1

Acquisition

For both dependent variables, performance during initial training (i.e., from Start-of-Practice to End-of-Practice) was assessed by 2[Group] \times 2[Time Point] ANOVAs with repeated measures on the last factor: (a) the *error score* (ES) was low from the beginning and decreased slightly, though not significantly, from the Start to End of Practice ($p = 0.128$). The groups did not differ ($p = 0.968$), nor was there a



Group \times Time Point interaction ($p = 0.773$); and (b) in contrast, the *sequence execution time* (TET) significantly decreased during acquisition ($F_{[\text{Time-Point}]}(1,22) = 330.44$, $p < 0.001$, $\eta_p^2 = 0.94$). However, there was neither a Group \times Time Point interaction ($p = 0.928$), nor was TET any different as a function of groups ($p = 0.686$). The TET performance did not level off at the end of acquisition: as became evident from a 2[Group] \times 3[Block] ANOVA, the TET significantly decreased even across the last three practice blocks ($F_{[\text{Block}]}(2,44) = 12.33$, $p < 0.001$, $\eta_p^2 = 0.36$), while neither the Block \times Group interaction nor the factor Group reached statistical significance ($p_{[\text{Group}]} = 0.688$; $p_{[\text{Block} \times \text{Group}]} = 0.17$). According to the respective pairwise comparisons, the TET significantly decreased from trial block to trial block until the end of practice ($p \leq 0.024$).

Retention

Again, for both dependent variables, performance during retention (i.e., from End-of-Practice to Retest) was assessed by $2[\text{Group}] \times 2[\text{Time Point}]$ ANOVAs with repeated measures on the last factor. The overall *error score* (ES) did not change further during retention ($p = 0.447$). The groups did not differ ($p = 0.384$), nor was there a significant Group \times Time point interaction ($p = 0.195$). Thus, throughout the 12-h retention interval, ES never deviated from its lower asymptote that was reached at the end of practice. TET was significantly shorter at Retest compared to End-of-Practice ($F_{(1,22)} = 28.38$, $p < 0.001$, $\eta_p^2 = 0.56$). There was a significant Group \times Time Point interaction ($F_{(1,22)} = 8.27$, $p = 0.009$, $\eta_p^2 = 0.27$), while the groups did not differ ($p = 0.866$). That is, the TET reduction at the Retest was significantly greater in the EM group than in the ME group. This result was confirmed by two paired t -tests, which were calculated for each group separately to compare the sequence execution time at the End-of-Practice and Retest. Here, a Bonferroni-corrected significance level of $p < 0.025$ was applied. It was shown that TET significantly decreased only during the sleep-filled retention interval in the EM group ($t_{(11)} = 5.16$, $p_{[\text{two-tailed}]} < 0.001$, $d = 1.49$) but not during the waking interval in the ME group ($t_{(11)} = 2.01$, $p_{[\text{two-tailed}]} = 0.069$, $d = 0.58$). The respective TET-change scores from the End-of-Practice to Retest amounted to 8.9% in the EM group ($p_{[\text{two-tailed}]} < 0.001$) and 4.1% in the ME group ($p_{[\text{two-tailed}]} = 0.114$).

Experiment 2

Acquisition

For both dependent variables, performance during initial training (i.e., from Start-of-Practice to End-of-Practice) was assessed again by $2[\text{Group}] \times 2[\text{Time Point}]$ ANOVAs with repeated measures on the last factor. The *error score* (ES) significantly decreased from the Start to End of Practice ($F_{[\text{Time-Point}](1,22)} = 9.41$, $p = 0.006$, $\eta_p^2 = 0.30$). The groups did not differ ($p = 0.280$), nor was there a Group \times Time Point interaction ($p = 0.650$). The number of ES per trial block dropped rapidly to a value < 1 after only 20 practice trials. After having reached that asymptote, the second half of the acquisition period ES remained fairly constant in both experimental groups. This was shown by a $2[\text{Group}] \times 5[\text{Block}]$ ANOVA calculated for blocks 6–10 ($p_{[\text{Block}]} = 0.533$; $p_{[\text{Group}]} = 0.085$; $p_{[\text{Block} \times \text{Group}]} = 0.628$). The *sequence execution time* (TET) also decreased significantly during acquisition ($F_{[\text{Time-Point}](1,22)} = 459.92$, $p < 0.001$, $\eta_p^2 = 0.95$), while again, neither the factor Group ($p = 0.724$), nor the interaction term ($p = 0.937$) reached statistical significance. Contrary to the error score, during the acquisition the TET never reached a performance asymptote. Instead, even across the last three practice blocks (i.e., blocks 8–10), the TET decreased significantly, as was shown by a $2[\text{Group}] \times 3[\text{Block}]$ ANOVA ($F_{[\text{Block}](2,44)} = 4.70$, $p = 0.014$, $\eta_p^2 = 0.17$; $p_{[\text{Group}]} = 0.756$; $p_{[\text{Block} \times \text{Group}]} = 0.692$).

Retention

Performance during retention (i.e., from End-of-Practice to Retest) for both dependent variables was assessed by

$2[\text{Group}] \times 2[\text{Time Point}]$ ANOVAs with repeated measures on the last factor. There was no further significant change in the *error score* (ES) during retention ($p = 0.098$), nor was there a significant group difference ($p = 0.273$) or Group \times Time Point interaction ($p = 0.733$). Additionally, the *sequence execution time* (TET) did not change throughout the 12-h retention interval but remained virtually constant ($p = 0.892$). Again, there were no group differences ($p = 0.681$) nor a Group \times Time Point interaction ($p = 0.856$). The respective TET-change scores from the End-of-Practice to Retest amounted to -0.4% in the EM group and in the ME group ($p_{[\text{two-tailed}]} \geq 0.77$).

Experiment 3

Acquisition

Again, for both dependent variables, performance during initial training (i.e., from Start-of-Practice to End-of-Practice) was assessed by $2[\text{Group}] \times 2[\text{Time Point}]$ ANOVAs with repeated measures on the last factor. The *error score* (ES) was almost negligible right from the beginning. There were no more than 2.5% ES per trial block throughout acquisition, and the ES did not change significantly from the Start to End of Practice ($p = 0.380$). The groups did not differ ($p = 0.873$), nor was there a Group \times Time Point interaction ($p = 0.640$). Contrary to this, the *sequence execution time* (TET) significantly differed as a function of groups ($F_{[\text{Group}](1,23)} = 6.01$, $p = 0.022$, $\eta_p^2 = 0.20$) and significantly decreased during acquisition ($F_{[\text{Time-Point}](1,23)} = 197.01$, $p < 0.001$, $\eta_p^2 = 0.89$). These performance improvements were about identical in both experimental groups, as was shown by a non-significant Group \times Time Point interaction ($p = 0.924$). Other than ES, the TET continuously decreased across the trial blocks and never reached an asymptote. In each experimental group, there was a significant decrease in TET even across the last three practice blocks (i.e., blocks 8–10), as was shown by a $2[\text{Group}] \times 3[\text{Block}]$ ANOVA ($F_{[\text{Block}](2,46)} = 3.29$, $p = 0.046$, $\eta_p^2 = 0.12$; $F_{[\text{Group}](1,23)} = 5.63$, $p = 0.026$, $\eta_p^2 = 0.19$; $p_{[\text{Block} \times \text{Group}]} = 0.488$).

Retention

Performance during retention (i.e., from End-of-Practice to Retest) for both dependent variables was assessed by $2[\text{Group}] \times 2[\text{Time Point}]$ ANOVAs with repeated measures on the last factor. There was no further significant change in the *error score* (ES) during retention ($p = 0.216$), nor was there a significant group difference ($p = 0.580$) or a Group \times Time Point interaction ($p = 0.839$). However, the *sequence execution time* (TET) significantly decreased again ($F_{(1,23)} = 14.80$, $p = 0.001$, $\eta_p^2 = 0.39$). The groups differed as before ($F_{(1,23)} = 5.96$, $p = 0.023$, $\eta_p^2 = 0.20$), but again, the interaction term failed to reach statistical significance ($p = 0.500$), indicating that TET-reduction throughout the 12-h retention interval was about the same in both experimental groups. The TET-offline improvements in both groups were quite similar and rather small. According to the paired t -tests calculated for each group separately, only those in the ME group were significant at a Bonferroni-corrected significance level

of $p < 0.025$ (EM group: $t_{(12)} = 2.52$, $p_{[\text{two-tailed}]} = 0.027$, $d = 0.71$; ME group: $t_{(11)} = 2.87$, $p_{[\text{two-tailed}]} = 0.015$, $d = 0.82$). The respective TET-change scores from the End-of-Practice to Retest amounted to 4.1% ($p_{[\text{two-tailed}]} = 0.025$) in the EM group and 4.9% ($p_{[\text{two-tailed}]} = 0.013$) in the ME group.

DISCUSSION

In the motor domain, task complexity is thought to be associated with the control requirements imposed on the human neuromotor system. In the initial stages of motor learning, task complexity is also thought to affect memory consolidation. A specific role in memory consolidation has been attributed to sleep-related processes enhancing the respective skill representation during retention in the absence of any further physical practice. These processes are thought to be causal for marked performance enhancements that have repeatedly been found at Retest only following sleep but not following waking (sleep-related offline learning). The objective of the three studies presented here was to identify whether task complexity affects sleep-related performance improvements in gross motor skills. Two different components of task complexity were addressed, namely, *sequence length* and *structural complexity*, with the latter being a measure of sequence regularity. For gross motor tasks, the sequences of unrestrained arm movements were employed. Each sequence was composed of a certain number of goal-directed reaching movements following a specific spatial pattern. These reaching movements are regarded as the basic units of behavior (i.e., sequence elements). In adults, they are largely pre-programmed before initiation (Gordon et al., 1994) and follow a fixed kinematic pattern (reaching synergy; see Konczak and Dichgans, 1997). During execution, the redundant degrees of freedom then are constrained by subconscious, lower-level control mechanisms to achieve optimal trajectory control and to compensate for any transient perturbations, thereby following a minimum cost principle (Cruse et al., 1993; Dounskaia and Wang, 2014). Initially these reaching movements are separate intentional acts. Organized in a defined ordinal succession they constitute the fundamental pattern of the respective gross motor skill.

The three arm movement sequences were modified according to each experiment's complexity requirements: (a) in *Experiment 1*, we used a 10-element sequence (nine different target locations) following a non-regular spatial pattern. The same sequence had been used in a previous study (Malangré and Blischke, 2016). In this study, sleep-related offline learning had been found when the participants were retested under free recall conditions. Therefore, sleep-related offline learning was also expected to occur in the present study, with stimulus information still available at retest; (b) in *Experiment 2*, only the *sequence length* was altered by reducing the number of sequence elements from 10 to 5. Although the structural task complexity remained unchanged, this shorter sequence was expected to be much simpler and easier to learn; and (c) in *Experiment 3*, the sequence length was the same as

in *Experiment 1*; only the *structural complexity* was reduced by arranging the sequence elements according to the regular geometric forms. The resulting pattern was expected to be learned and remembered much easier than the one employed in *Experiment 1*. In reaching movements of this type, movement time depends on their ID (Fitts, 1954). To avoid precision requirements confounding the total sequence execution time (TET), the mean ID was kept about the same in the respective arm movement sequences.

In all three experiments, the participants initially practiced criterion tasks either in the morning (the ME group) or in the evening (the EM group) and were retested 12 h later. Thus, during retention, the EM groups were afforded a night of sleep while the ME groups had to stay awake. The dependent variables were performance error (i.e., number of ES per trial block) and TET per sequence (correct sequences only). Thus, performance increases as the error score and/or execution time decreases. Note that TET reduction is proportional to an increase in sequence execution speed. The amount of practice and trial number at the Retests were identical in all experiments. In all three experiments, the error score was low from the beginning and remained stable until end of practice without any between-group differences. Across the delay period, there were no significant differences in performance error within or between groups. This indicates that no loss of accuracy accompanied improvements in the sequence execution speed. Instead, all performance differences/changes were exclusively reflected by the TET. The TET significantly decreased during acquisition in all three experiments. At Retest, in *Experiment 1*, the TET significantly decreased again following sleep but not after waking. Any such offline improvements were completely eliminated, however, when the sequence length was reduced to only five elements in *Experiment 2*. In *Experiment 3*, small but significant offline improvements in terms of TET reduction occurred again, but this time, these improvements were of the same magnitude following sleep or waking. Moreover, offline improvements in *Experiment 3* were only half the size of those following sleep in *Experiment 1*.

According to these results, sleep-related offline learning appears to be associated with a sufficient amount of task complexity. In contrast, when the task complexity is sufficiently reduced by either reducing the sequence length or increasing the regularities underlying sequence production, sleep-related performance enhancements are not observed any longer. However, before coming to a final conclusion, several methodological issues first need to be discussed. With respect to motor sequence learning, it has recently been argued that there is little evidence for a performance gain that can be attributed to sleep when confounding variables are factored out that are independent of any possible sleep consolidation effect (e.g., Pan and Rickard, 2015). Here, confounds due to online learning during retests, data averaging and reactive inhibition have been found to be of prominent significance. We will briefly address these aspects in the following section.

In a wide variety of task domains, performance improvement both within and between sessions is known to follow a smooth, monotonically decreasing curve (Newell and Rosenbloom, 1981). Accordingly, Pan and Rickard (2015) have proposed to solve the problem of an *online learning* confound by fitting an appropriate empirical function to each subject's practice data. The gain score analysis then can be based on the difference between the predicted retest performance (based on the extrapolation of the practice data fit) and the observed performance on the retest blocks. Power law functions ($f(x) = ax^{-\alpha}$) have been used previously to estimate test performance in previous studies concerning sleep-related offline learning (e.g., Adijapha et al., 2014; Malangré et al., 2014). The assumption of performance improvements closely resembling a power law function holds true also with respect to our own data. This was clearly shown in a study with eight subjects that extensively practiced the movement sequence used in our present Experiment 1 for 600 trials, which was distributed over 3 days (Schmitz and Waßmuth, 2013; unpublished data). The TET continuously decreased following a power function and approached an asymptote of approximately 5.7 s on average only after 550 trials (Supplementary Figure S2). We therefore employed curve fitting to the TET data also in the present case. According to this approach (see Supplementary data for details), sleep-related offline learning appears to be confirmed in Experiment 1. Here, the observed retest performance is significantly superior to the predicted performance in the EM group. Conversely, in the ME group, the respective TET data account for continued online learning following the wake retention interval. That is, only those individuals experiencing a night of sleep revealed improvements in a sequence execution speed at retest beyond what was predicted by individual learning functions derived from their initial practice performance. In Experiment 2, the observed TET-performance at retest was slightly worse than the predicted one. This may be due to a pronounced warm-up decrement evident in both experimental groups in the first retest trial block (see Supplementary Figure S1). In Experiment 3, continued online learning accounts for the slight performance increments that were found at retest. In both of the experimental groups the observed performance measures and the predicted performance measures are about the same.

Data averaging procedures are commonly used to assess post-sleep gain throughout the motor consolidation literature because they effectively compensate local performance fluctuations. Therefore, we also averaged performance data at each time point over three trial blocks. However, with respect to testing the sleep-related enhancement hypothesis, data averaging can also profoundly bias results. Averaging over the steeper section of the performance curve early in practice may exacerbate online learning confounds. Moreover, averaging can also be highly sensitive to transient performance patterns, such as warm-up effects on initial test blocks. According to Pan and Rickard (2015), addressing such warm-up effects by just eliminating data from the first test block prior to the calculation of change scores would be ill-advised because it may even exacerbate the

online learning confound discussed above. Instead, these authors recommend curve fitting as the one approach that can fully resolve the respective confound due to data averaging.

It is well known that massed practiced conditions can result in an accumulating decline in performance, which is often associated with scalloped *reactive inhibition* effects (for an overview, see Pan and Rickard, 2015). Therefore, any offline gain observed following a delay period at retest may result from the differences in magnitude of reactive inhibition at the end of practice vs. the beginning of the test and may not require a sleep consolidation interpretation. However, reactive inhibition effects can resolve relatively quickly when brief breaks are inserted between blocks of training or when post-training performance is assessed only a few minutes after the end of practice. In our present data, we did not find any evidence of scalloped reactive inhibition. Rather, within practice blocks, early and late trials do not statistically differ, while across practice blocks, performance increases even at the end of practice. This may be due to the small number of trials per block as well as the fact that our participants were instructed to start a new trial within a block only after they had signaled that they were ready to do so. Moreover, in the already mentioned study by Malangré and Blischke (2016) with the same arm movement sequence employed as in our present Experiment 1, post-training measures were assessed under free recall conditions 15 min after practice. After that time, according to the material reviewed by Pan and Rickard (2015), any reactive inhibition effects would have entirely resolved already. Nevertheless, in that study, performance significantly *increased* once more relative to this early post-training test by approximately 10% when the subjects were retested again following sleep and appeared to be stabilized after waking. These results also are in clear support of the sleep-related performance enhancement hypothesis. Thus, all in all, we do not think that reactive inhibition really affected our present data.

There are some other factors that have also been considered to be possible offline learning confounds. According to Pan and Rickard (2015), the *time of training* is not critical to the observed gain score, specifically when the effects of morning vs. evening training times are compared. To fully equate training and testing sessions with respect to physiological circadian and homeostatic factors, however, it appears prudent to use a 24-h delay design. This is what we did in two previous studies using the same task as in Experiment 1 (Malangré and Blischke, 2016) or a task similar to that (Malangré et al., 2014). Independent of time of training, in both of these studies, significant performance improvements were found following sleep but not after waking. We therefore thought it sufficient to limit our present experimental design to a 12-h retention period. Within this time frame, the findings from our present Experiment 1 replicate the previous results well. Conversely, the *time of testing* may have a larger influence on the post-delay gain. For instance, Nettersheim et al. (2015) recently presented data suggesting a transient performance boost in finger-tapping skills 30 min after the end of practice. According to their findings, sleeping immediately afterwards stabilized performance

at this level but did not enhance it. However, sleep applied considerably later enhanced performance relative to that assessed at the end of practice only to the “transient boost” level. These and other findings (e.g., Landry et al., 2016) call into question some concepts concerning sleep-related mechanisms of motor memory consolidation that are presently debated. For the time being, however, generalizing these findings across motor task domains would probably be precipitous. In this respect, future work will certainly be needed.

So far, we have dealt with critical issues basically related to the chosen study design. However, another point at issue relates to the *type of motor sequence representation*, which is supposed to be enhanced offline while subjects are asleep. According to the widely accepted model initially proposed by Hikosaka et al. (1999), the acquisition of sequential behaviors resides in the interaction between different neural networks that would encode the same motor sequence in two different coordinate systems (i.e., spatial and motor). One memory component is created rapidly early during training and is thought to incorporate allocentric (spatial) coordinates (e.g., spatial locations of end effectors and/or sequential target positions) and to constitute an abstract effector-independent representation of a series of movements that need to be executed in an external frame of reference. The respective representational code relies on attention, explicit knowledge and working memory. The other memory component develops more slowly and is supposed to be mediated through egocentric (motor) coordinates (e.g., sequence of activation patterns of the agonist/antagonist muscles and/or achieved joint angles) and thus should constitute an effector-dependent, movement-based skill realized in an internal frame of reference. This representation relies on implicit knowledge and does not require attention or working memory. Hikosaka's model has also been applied to research motor memory consolidation. There, it has been shown that *sleep* specifically favors enhancement of the *extrinsic* (spatial) sequence representation, while consolidation of the respective *intrinsic* (motor) representation was *not* modulated by the sleep/wake condition (Witt et al., 2010; Albouy et al., 2013). According to this model, and considering that after only 100 sequence execution trials our subjects still had only little practice, we conjecture performance improvements observed in our experiments to be mainly due to enhancement of the abstract spatial sequence representation and its declarative knowledge base.

Still another caveat might be raised considering our chosen main criterion variable, namely, *TET*. As long as participants have not yet fully memorized the spatial sequence pattern specifying the target locations on the pegboard, they need to visually register each new target stimulus on the computer screen and then find its counterpart on the pegboard. Thus, it could be argued that our participants' performance improvements (i.e., the *TET*-reduction) substantially resulted from improving their gaze behavior. However, in our opinion this does not call into question the validity of *TET* (i.e., sequence execution time) as a global performance measure. The visual target location is always an inherent component of any reaching movement executed in Euclidean space. The respective saccades and

reaching movements are usually initiated simultaneously, with the saccades being terminated earlier than the limb movement (Jeannerod, 1997). The visual target location must therefore be regarded an integral part of the sequence execution skill at any stage of expertise. Early on, however, the subjects are still in the process of acquiring a cognitive, abstract spatial sequence representation (Hikosaka et al., 1999). During that stage, the gaze moves back and forth between the computer screen and work space, and a visual search will necessarily add to *TET* as the overall performance variable. However, once a sequence pattern has been learned, subjects do not need to rely on visual stimulus information any longer. This has been demonstrated in the previously mentioned study by Malangré and Blischke (2016), where sleep-related offline improvements have been shown under free recall conditions.

In that previous study, the criterion task was either identical to (Experiment 1) or more difficult than the tasks used in our present study (Experiments 2 and 3). At the same time, however, the study by Malangré and Blischke (2016) involved 20% more practice trials (i.e., a total of 120 trials) than our present experiments, and even then, six out of a total of 24 participants eventually had to be excluded from the final analysis for not having fully memorized the sequence. We therefore assume that in our present study, at least in Experiment 1, a considerable albeit unknown number of participants had not yet fully learned the sequence prior to Retest. These subjects are likely to have relied on visual checks and cuing again when being retested. The time attributed to this gaze behavior therefore could have obscured the true effect of sleep-related memory consolidation to some extent. Moreover, suchlike incomplete declarative learning during acquisition in all probability was more pronounced in Experiment 1 than in the two other experiments, because their tasks were easier and could be memorized much sooner. We therefore have to acknowledge that different levels of explicit sequence knowledge in the different experiments might also have contributed to the differential sleep-dependent improvement. These limitations call for a careful interpretation of our present results, and certainly more research is needed in order to corroborate our present findings.

All this considered, we still believe that our present data provide some new evidence on the relationship between task complexity and motor memory consolidation. More specifically, for the first time, we have shown that the number of elements and structural complexity *independently* affect sleep-related offline performance improvements in a sequential motor task under cued recall conditions. That is, reducing structural complexity in a ten-element sequence indeed had the same effect as reducing the sequence length from 10 to 5 elements, with structural complexity remaining invariably high. In both cases, sleep-related performance enhancements that are evident in a sufficiently complex arm movement sequence were not observed any longer. Thus, by dissociating different complexity components and by incorporating gross motor tasks, we successfully extended earlier work by Kuriyama et al. (2004).

The question now arises as to how our observations might be interpreted in light of theoretical considerations regarding processes of sleep-related motor memory consolidation. One possible answer might be found in relation to processing demands associated with sequence acquisition, the participants' working memory capacity, and subsequent processes of sleep-related memory enhancement. Sequences that are easy to remember might not profit as much from active system consolidation processes as sequences that clearly exceed working memory capacity. Sequences of up to five independent elements (as in Experiment 2) are well within adult subjects' memory span (Mathy and Feldman, 2012) and thus are easy to remember during acquisition. This might have limited the performance-enhancing effect of subsequent consolidation processes. On the other hand, while longer sequences are more difficult to remember, the simplicity of patterns within a sequence yields an independent influence: a sequence of given length is more difficult to remember as the number of distinct subsets (chunks) increases. Conversely, fewer and larger subsets within a sequence reduce its structural complexity (and increase its formal "compressibility"; see Mathy and Feldman, 2012). Accordingly, the complexity of a sequence expresses how much memory space is required to encode it and determines the success of memory processes. Again, if a sequence representation is already well-established during initial practice due to its low structural complexity, subsequent consolidation processes may not be able to induce any further performance improvements. Specifically, with respect to Experiment 3, we conjectured that tracing familiar geometric forms in the course of sequence execution would promote the use of cognitive strategies, which in turn could serve as binding rules to integrate the individual task components into a meaningful "gestalt". The respective geometric arrangement of a set of separate reaching movements would then immediately be conceptualized as a unified representation (Franz et al., 2001; Swinnen and Wenderoth, 2004) and effectively reduce the structural task complexity. As a consequence, we believe that a sufficiently stable sequence representation was already established during practice and that subsequent memory consolidation affected the performance at retest to a lesser degree. Of course, all these considerations are only preliminary at present and need to be carefully scrutinized in the course of future research.

REFERENCES

- Adi-Japha, E., Badir, R., Dorfberger, S., and Karni, A. (2014). A matter of time: rapid motor memory stabilization in childhood. *Dev. Sci.* 17, 424–433. doi: 10.1111/desc.12132
- Albouy, G., Fogel, S., King, B. R., Laventure, S., Benali, H., Karni, A., et al. (2015). Maintaining vs. enhancing motor sequence memories: respective roles of striatal and hippocampal systems. *Neuroimage* 108, 423–434. doi: 10.1016/j.neuroimage.2014.12.049
- Albouy, G., Fogel, S., Pottiez, H., Nguyen, V. A., Ray, L., Lungu, O., et al. (2013). Daytime sleep enhances consolidation of the spatial but not motoric representation of motor sequence memory. *PLoS One* 8:e52805. doi: 10.1371/journal.pone.0052805

STUDY LIMITATIONS

There are also some limitations to our study: (a) one is the lack of polysomnographic data. In future studies, it would be interesting to see the extent to which the impact of motor task complexity on offline learning in a gross motor task is also systematically reflected in polysomnographic measures such as, for example, sleep-spindle density; (b) another limitation arises from the fact that we did not assess our participants' individual working memory capacity. In the context of the above considerations, it appears reasonable to expect task complexity and working memory capacity to interact in their impact on motor memory consolidation at the individual level as well; (c) we also did not systematically assess our participants' cognitive retrieval strategies. This could have contributed to a better understanding of the declarative knowledge underlying successful sequence retrieval; (d) furthermore, we experimentally addressed only two components of motor task complexity so far (i.e., the number of sequence elements and structural complexity). Incorporating components such as, for example, the number of end-effectors or inter-limb coordination, was beyond the scope of our present study. However, these aspects are specifically relevant to tasks in applied areas such as sports and vocational training, stroke rehabilitation and physical therapy; and (e) last, in future studies, assessment of gaze behavior and kinematic data could elucidate how different components involved in gross motor sequence execution might contribute differently to performance improvements due to practice and subsequent memory consolidation.

AUTHOR CONTRIBUTIONS

Both authors contributed extensively and equally to the work presented in this manuscript; discussed and commented on the manuscript at all stages. KB and AM developed the research topic and designed the experiments; prepared, analyzed data and discussed results. AM supervised data collection. KB wrote the article.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fnhum.2017.00374/full#supplementary-material>

- Al-Sharman, A., and Siengsukon, C. F. (2013). Sleep enhances learning of a functional motor task in young adults. *Phys. Ther.* 93, 1625–1635. doi: 10.2522/ptj.20120502
- Backhaus, W., Braaß, H., Renné, T., Krüger, C., Gerloff, C., and Hummel, F. C. (2016). Daytime sleep has no effect on the time course of motor sequence and visuomotor adaptation learning. *Neurobiol. Learn. Mem.* 131, 147–154. doi: 10.1016/j.nlm.2016.03.017
- Born, J., and Wilhelm, I. (2012). System consolidation of memory during sleep. *Psychol. Res.* 76, 192–203. doi: 10.1007/s00426-011-0335-6
- Boyle, J. B., and Shea, C. H. (2011). Wrist and arm movements of varying difficulties. *Acta Psychol.* 137, 382–396. doi: 10.1016/j.actpsy.2011.04.008

- Cohen, D. A., Pascual-Leone, A., Press, D. Z., and Robertson, E. M. (2005). Off-line learning of motor skill memory: a double dissociation of goal and movement. *Proc. Natl. Acad. Sci. U S A* 102, 18237–18241. doi: 10.1073/pnas.0506072102
- Cruse, H., Brüewer, M., and Dean, J. (1993). Control of three- and four-joint arm movement: strategies for a manipulator with redundant degrees of freedom. *J. Mot. Behav.* 25, 131–139. doi: 10.1080/00222895.1993.9942044
- Dounskaia, N., and Wang, W. (2014). A preferred pattern of joint coordination during arm movements with redundant degrees of freedom. *J. Neurophysiol.* 112, 1040–1053. doi: 10.1152/jn.00082.2014
- Doyon, J., Korman, M., Morin, A., Dostie, V., Tahar, A., Benali, H., et al. (2009). Contribution of night and day sleep vs. simple passage of time to the consolidation of motor sequence and visuomotor adaptation learning. *Exp. Brain Res.* 195, 15–26. doi: 10.1007/s00221-009-1748-y
- Fischer, S., Nitschke, M. F., Melchert, U. H., Erdmann, C., and Born, J. (2005). Motor memory consolidation in sleep shapes more effective neural representations. *J. Neurosci.* 25, 11248–11255. doi: 10.1523/JNEUROSCI.1743-05.2005
- Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *J. Exp. Psychol.* 47, 381–391. doi: 10.1037/h0055392
- Franz, E. A., Zelaznik, H. N., Swinnen, S. S., and Walters, C. (2001). Spatial conceptual influences on the coordination of bimanual actions: when a dual task becomes a single task. *J. Mot. Behav.* 33, 103–112. doi: 10.1080/00222890109601906
- Genzel, L., Kroes, M. C. W., Dresler, M., and Battaglia, F. P. (2014). Light sleep versus slow wave sleep in memory consolidation: a question of global versus local processes? *Trends Neurosci.* 37, 10–19. doi: 10.1016/j.tins.2013.10.002
- Genzel, L., Quack, A., Jäger, E., Konrad, B., Steiger, A., and Dresler, M. (2012). Complex motor sequence skills profit from sleep. *Neuropsychobiology* 66, 237–243. doi: 10.1159/000341878
- Goertelmeyer, R. (1986). “Schlaf-Fragebogen A und B (Sf-A, Sf-B),” in *Collegium Internationale Psychiatriae Scalarum*, 3rd Edn. ed. International Skalen für Psychiatrie (Weinheim: Beltz).
- Gordon, J., Ghilardi, M. F., and Ghez, C. (1994). Accuracy of planar reaching movements: I. Independence of direction and extent variability. *Exp. Brain Res.* 99, 97–111. doi: 10.1007/bf00241415
- Gudberg, C., Wulff, K., and Johansen-Berg, H. (2015). Sleep-related motor memory consolidation in older adults depends on task demands. *Neurobiol. Aging* 36, 1409–1416. doi: 10.1016/j.neurobiolaging.2014.12.014
- Hikosaka, O., Nakahara, K., Rand, M. K., Sakai, K., Lu, X., Nakamura, K., et al. (1999). Parallel neural networks for learning sequential procedures. *Trends Neurosci.* 22, 464–471. doi: 10.1016/s0166-2236(99)01439-3
- Hoedlmoser, K., Birklbauer, J., Schabus, M., Eibenberger, P., Rigler, S., and Mueller, E. (2015). The impact of diurnal sleep on the consolidation of a complex gross motor adaptation task. *J. Sleep Res.* 24, 100–109. doi: 10.1111/jsr.12207
- Jeannerod, M. (1997). *The Cognitive Neuroscience of Action*. Cambridge, MA: Blackwell.
- Kempler, L., and Richmond, J. L. (2012). Effect of sleep on gross motor memory. *Memory* 20, 907–914. doi: 10.1080/09658211.2012.711837
- Konczak, J., and Dichgans, J. (1997). The development towards stereotype arm kinematics during reaching in the first three years of life. *Exp. Brain Res.* 117, 346–354. doi: 10.1007/s002210050228
- Kuriyama, K., Stickgold, R., and Walker, M. P. (2004). Sleep-dependent learning and motor-skill complexity. *Learn. Mem.* 11, 705–713. doi: 10.1101/lm.76304
- Landry, S., Anderson, C., and Conduit, R. (2016). The effects of sleep, wake activity and time-on-task on offline motor sequence learning. *Neurobiol. Learn. Mem.* 127, 56–63. doi: 10.1016/j.nlm.2015.11.009
- Malangré, A., and Blischke, K. (2016). Sleep-related offline improvements in gross motor task performance occur under free recall requirements. *Front. Hum. Neurosci.* 10:134. doi: 10.3389/fnhum.2016.00134
- Malangré, A., Leinen, P., and Blischke, K. (2014). Sleep-related offline learning in a complex arm movement sequence. *J. Hum. Kinet.* 40, 7–20. doi: 10.2478/hukin-2014-0002
- Mathy, F., and Feldman, J. (2012). What’s magic about magic numbers? Chunking and data compression in short-term memory. *Cognition* 122, 346–362. doi: 10.1016/j.cognition.2011.11.003
- Morita, Y., Ogawa, K., and Uchida, S. (2012). The effect of a daytime 2-hour nap on complex motor skill learning. *Sleep Biol. Rhythms* 10, 302–309. doi: 10.1111/j.1479-8425.2012.00576.x
- Nemeth, D., Janacek, K., Londe, Z., Ullman, M. T., Howard, D. V., and Howard, J. H. Jr. (2010). Sleep has no critical role in implicit motor sequence learning in young and old adults. *Exp. Brain Res.* 201, 351–358. doi: 10.1007/s00221-009-2024-x
- Nettersheim, A., Hallschmidt, M., Born, J., and Diekelmann, S. (2015). The role of sleep in motor sequence consolidation: stabilization rather than enhancement. *J. Neurosci.* 35, 6696–6702. doi: 10.1523/JNEUROSCI.1236-14.2015
- Newell, A., and Rosenbloom, P. S. (1981). “Mechanisms of skill acquisition and the law of practice,” in *Cognitive Skills and Their Acquisition*, ed. J. R. Anderson (Hillsdale, NJ: Erlbaum), 1–55.
- Pan, S. C., and Rickard, T. C. (2015). Sleep and motor learning: is there room for consolidation? *Psychol. Bull.* 141, 812–834. doi: 10.1037/bul0000009
- Robertson, E. M., Pascual-Leone, A., and Press, D. Z. (2004). Awareness modifies the skill-learning benefits of sleep. *Curr. Biol.* 14, 208–212. doi: 10.1016/s0960-9822(04)00039-9
- Schmitz, L., and Waßmuth, N. (2013). *Zum Einfluß der Übungsrate auf Ausführungsgeschwindigkeit und Reproduktionssicherheit einer komplexen Armbewegungsfolge—Eine Laborstudie*. Unpublished Bachelors’ Thesis; Department of Sport Science, Saarland University, Germany.
- Smith, C. (1995). Sleep states and memory processes. *Behav. Brain Res.* 69, 137–145. doi: 10.1016/0166-4328(95)00024-n
- Song, S., Howard, J. H. Jr., and Howard, D. V. (2007). Sleep does not benefit probabilistic motor sequence learning. *J. Neurosci.* 27, 12475–12483. doi: 10.1523/JNEUROSCI.2062-07.2007
- Stickgold, R., Hobson, J. A., Fosse, R., and Fosse, M. (2001). Sleep, learning, and dreams: off-line memory reprocessing. *Science* 294, 1052–1057. doi: 10.1126/science.1063530
- Swinnen, S. P., and Wenderoth, N. (2004). Two hands, one brain: cognitive neuroscience of bimanual skill. *Trends Cogn. Sci.* 8, 18–25. doi: 10.1016/j.tics.2003.10.017
- Walker, M. P. (2005). A refined model of sleep and the time course of memory formation. *Behav. Brain Sci.* 28, 51–64; discussion 64–104. doi: 10.1017/s0140525x05000026
- Walker, M. P., Brakefield, T., Morgan, A., Hobson, J., and Stickgold, R. (2002). Practice with sleep makes perfect: sleep-dependent motor skill learning. *Neuron* 35, 205–211. doi: 10.1016/S0896-6273(02)00746-8
- Witt, K., Margraf, N., Bieber, C., Born, J., and Deuschl, G. (2010). Sleep consolidates the effector-independent representation of a motor skill. *Neuroscience* 171, 227–234. doi: 10.1016/j.neuroscience.2010.07.062

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Blischke and Malangré. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.